National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

UNIVERSITY OF ALBERTA

MOLECULAR STUDIES OF PROTEIN-DNA INTERACTIONS AT THE ORIGIN OF TRANSFER OF CONJUGATIVE PLASMIDS

BY

LAURA DI LAURENZIO ©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
AND RESEARCH IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY

DEPARTMENT OF BIOCHEMISTRY

EDMONTON, ALBERTA

SPRING, 1992

National Library of Canada
Canadian Theses Service
Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Service des thèses canadiennes

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-73275-X

Canada

ACTIVITY REPORT

TRANSMISSION OK

TRANSACTION #	1082
CONNECTION TEL	031 226 3803
CONNECTION ID	BSP EDINBURGH
START TIME	04/08 14:12
USAGE TIME	01'05
PAGES	2

FAX MESSAGE

DATE: APRIL 08 1992

TO: Ms Helen Ure
Executive Editor
Molecular Microbiology

FAX #: 044 ~~0~~31-226-3803

FROM: LAURA DI LAURENZIO

DEPARTMENT OF MICROBIOLOGY
UNIVERSITY OF ALBERTA
EDMONTON, AB T6G 2E9
CANADA
FAX # (403) 492-2216

MESSAGE: Dear Ms Ure

I would greatly appreciate if you could answer this letter as soon as possible.

Thanking in advance,

Sincerely yours,

Laura Di Laurenzio

TOTAL NUMBER OF PAGES 2 (INCLUDING COVER PAGE)

PLEASE CALL (403) 492-7348 IF THERE IS A PROBLEM WITH ANY PART OF THIS MESSAGE.

University of Alberta
Edmonton

Department of Microbiology

Canada T6G 2E9

M330 Biological Sciences Building, Telephone (403) 492-3277
Fax (403) 492-2216

April 8, 1992

Helen Ure
Molecular Microbiology
Blackwell Scientific Publications Ltd
23 Ainslie Place
Edinburgh EH3 6AJ
Scotland

Dear Ms Ure,

I am a graduate student at the University of Alberta and I am in the process of writing my PhD thesis. I would like to receive written permission to use the data and the figures published in the following article: L. Di Laurenzio, L. S. Frost, B. B. Finlay and W. Paranchych (1991), Characterization of the *oriT* region of the IncFV plasmid pED208. Molecular Microbiology, **5**, 1779-1790.

Looking forward to hearing from you soon,

Sincerely yours,

Laura Di Laurenzio

Laura Di Laurenzio
graduate student

Laura Di Laurenzio
M334 Bio Sci Bldg
University of Alberta
Edmonton T6G 2E9
Alberta Canada

BLACKWELL SCIENTIFIC PUBLICATIONS LTD

Nigel Blackwell *Chairman*
Robert Campbell *Managing Director*
Keith Bowker
Jonathan Conibear
Oluf Møller
John Robson
Peter Saugman
Martin Wilkinson

OSNEY MEAD
OXFORD
OX2 0EL

Telephone
Oxford (0865) 240201
International
+ 44 865 240201
DIALCOM: 79: BSP001
Fax (GRP 3)
(0865) 721205
Telegrams
Research, Oxford
Telex
83355 Medbok G

Ref:OP/TJC/CJR2

21 April 1992

Fax:0101 403 492-2216

Laura Di Laurenzio
Department of Microbiology
University of Alberta
Edmonton, AB T6G 2E9
Canada

Dear Ms Laurenzio

Re:Molecular Microbiology

Thank you for your fax message dated 8th April 1992, which was sent to our Edinburgh office, which we received on the 16th April 1992.

We are happy to grant you permission to reproduce the figure and illustrations from your article which appeared on pages 1779-1790 of Vol 5:7, subject to the usual acknowledgements of source.

Yours sincerely

T. Clements

Tracey Clements
Permissions Assistant

Registered Office: Osney Mead, Oxford, OX2 0EL Registration No: 349566 England
Also at 25 John Street, London, WC1N 2BL; 23 Ainslie Place, Edinburgh, EH3 6AJ;
3 Cambridge Center, Suite 208, Cambridge, Massachusetts 02142, U.S.A.; 54 University Street, Carlton, Victoria 3054, Australia
Offices also in Paris, Berlin and Vienna

TEXAS A&M
UNIVERSITY
HEALTH
SCIENCE
CENTER

College of Medicine

Department of Medical Microbiology & Immunology

March 6, 1992

Fax 403-492-2216

Laura Di Laurenzio
Dept. of Microbiology
University of Alberta
Edmonton, AB T6G 2E9
Canada

Dear Laura,

You are welcome to use the F transfer operon map figure that I distributed at the Banff meeting in your thesis.

Sincerely yours,

Karin Ippen-Ihler

Page 5
Fig 1.1

407 Reynolds Medical Building • College Station, Texas 77843-1114 • 409/845-1313 • FAX 409/845-3479
[illegible] Community Hospital

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Laura Di Laurenzio

TITLE OF THESIS: Molecular Studies of protein-DNA Interactions at the Origin of Transfer of Conjugative Plasmids

DEGREE: Doctor of Philosophy

YEAR THIS DEGREE IS GRANTED: Spring, 1992

Permission is hereby granted to the UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purpose only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

Laura D Laurenzio

Laura Di Laurenzio
c/o Dept. of Microbiology
University of Alberta
Edmonton, Alberta
T6G 2E9

DATE: APRIL 23, 1992

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled Molecular Studies of Protein-DNA Interactions at the Origin of Transfer of Conjugative Plasmids submitted by Laura Di Laurenzio in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Wm Paranchych
Dr. William Paranchych, Supervisor

WR Addison
Dr. William R. Addison

A. Richard Morgan
Dr. A. Richard Morgan

Susan M Rosenberg
Dr. Susan M. Rosenberg

DG Scraba
Dr. Douglas G. Scraba

Walter B Dempsey
Dr. Walter B. Dempsey, External Examiner

DATE: APRIL 24, 1992

Ai miei genitori Camillo ed Antonietta

ai miei fratelli e sorelle Luca, Andrea, Donatella e Cristina

per l'appoggio morale e l'affetto mostratomi

durante questi anni in Canada

A mio nonno Antonio per la lezione di vita

ABSTRACT

DNA sequence analysis of a 2.2 kb *Eco*RI-*Hin*dIII fragment from pED208, the derepressed form for pilus formation of the IncFV plasmid F_0*lac*, reveals sequences homologous to the origin of transfer or *oriT* region, *traM* and *traJ* genes of other IncF plasmids. The TraM protein was purified and immunoblots of fractionated cells containing pED208 or F_0*lac* showed that TraM is predominately in the cytoplasm. The TraM protein was found to bind to three large motifs in the *oriT* region: *sbmA*, an inverted repeat, *sbmB*, two direct repeats and *sbmC*, the *traM* promoter region. These three regions contain the sequence GANTC 16 times, which were spaced 11-12 bp (or multiples thereof) apart, suggesting that TraM protein binds in a complex manner over this entire region.

To determine the minimal binding site recognized by the TraM protein a series of oligonucleotides were synthesized. These oligonucleotides were used in electrophoretic mobility retardation assays as substrates for binding. TraM binds to the single- and double-stranded form of an oligonucleotide containing two GATTC sequences that are 11 bp apart but not to a single *Hin*fI site sequence. Circular dichroism spectra of TraM incubated with either the single- or the double-stranded oligonucleotide showed interaction between the protein and the DNA.

The gene encoding the TraM protein of the conjugative plasmid F was cloned, overexpressed and the gene product was purified. The TraM protein was found in the cytoplasm of cells carrying the F plasmid. DNase I footprinting experiments show that the purified protein protects three regions in the F *oriT* locus with different affinity for

the two strands of DNA. Within the protected regions a 15 nucleotide motif was identified that represents the DNA binding site. The TraM protein was also found to bind to a sequence in the *oriT* region of the non-conjugative plasmid ColE1. This resembles the three binding sites in the F *oriT* region.

The physical properties of the TraM proteins from F and pED208 were characterized by means of sedimentation equilibrium, far ultraviolet circular dichroism (CD), sucrose gradients and immunoblotting. The two proteins share a high degree of amino acid sequence similarity and quaternary structure, as well as similar secondary structures characterized by high α-helical content. Both proteins have been shown to bind DNA, but they do not have a sequence that resembles the helix-turn-helix motif typical of prokaryotic DNA-binding proteins.

The interaction between the Integration Host Factor protein (IHF) of *E. coli* and the *oriT* region of the conjugative plasmid pED208 was analysed using electrophoretic mobility retardation assays and DNase I protection experiments. In addition, the protein-DNA complexes were visualised by electron microscopy. IHF binds to three sites in the *oriT* region with different affinity, with the first site being the closest to the *oriT* site. The other two sites could only be demonstrated to interact with IHF if potassium glutamate replaced potassium chloride in the incubation mixtures. The position of the high affinity site is consistent with those already found for the F and R100 *oriT* regions; this supports the postulated functional role of IHF in the process involving DNA nicking and the initiation of DNA transfer.

ACKNOWLEDGMENTS

It is a great pleasure to acknowledge the people who have been instrumental to the completion of this thesis.

First of all, I would like to thank my supervisor, Dr. William Paranchych for his encouragement and support throughout the course of my studies and for always being available to listen to me with infinite patience. His commitment to the progress of Science and his diplomatic skills are qualities I should always strive for.

I would also like to thank the members of his laboratory both in the Biochemistry and in the Microbiology Departments who contributed so much to my scientific as well as social life.

My fellow graduate students: Dr. Britt Pasloske, Qiming Sun, the first friend I made in Canada, Stuart Lee, Dawn MacDonald and Barb Conway.

The technicians: Kathy Volpel, her good nature and her smile made the life of this worrywart graduate student much more pleasant, Stasia Paczkowski, the endless number of protein gels she poured for me deserves my eternal gratitude, Linda Glasier and Neil Ellert, their help in antisera preparation saved me from the anguish and the guilty feelings I have when poking another living organism.

The research associates; Dr. Parimi Sastry, his expertise in protein chemistry was greatly appreciated in my first attempts at open column chromatography, Dr. Mark Joffe and Dr. Dick Zoutman, their vision of Science as clinicians and their keen efforts in learning the ropes of molecular biology were truly amazing, and the recent addition to the laboratory: Dr. Pat Lane-Bell, Dr. Mark Farinha and Dr. Heinz Hahn.

Finally, Dr Laura Frost deserves a special thank you for kindly accepting the task of being my unofficial supervisor. Both as research associate in Dr. Paranchych's laboratory and now as professor in Microbiology, Laura has always

been there to listen to, advise, guide and cheer me up in my scientific endeavors. Her assistance and help in writing this thesis are sincerely appreciated. Her love for Science and for her family are lessons of life that I will treasure forever. I would also like to thank Laura for being such an inexhaustible organizer of trips to the Rockies.

I am grateful to the members of her laboratory: Sonya Penfold, Tim van Biesen, Karen Anthony, John Simon and Sheryl Ronald for adopting me as their "senior graduate student". In particular I would like to thank Sonya for patiently "combing" the thesis in search of spelling mistakes.

In the course of the past five years several professors have been part of my committee, I wish to thank them all for being always helpful and supportive. In particular, I would like to thank: Dr. Larry Smillie for extending his use of his laboratory facilities for protein purification, amino acid sequencing and amino acid analysis, Dr. Bill Addison for the lessons in "DNase I footprinting" and " how to handle radioactivity without the use of gloves", Dr. Doug Scraba for sharing his expertise in electron microscopy and for the use of the digitizer, Dr. Susan Rosenberg who gracefully accepted to be the last minute addition to my committee and read the thesis on short notice, Dr. Dick Morgan for the many helpful discussions on DNA supercoiling and the use of his fluorimeter, but above all I would like to thank him for the encouragement during the last hectic weeks of preparation of this thesis.

I wish also to thank Dr. Cyril Kay for the use of his laboratory facilities for the physico-chemical characterization of the proteins and for the critical review of Chapters III and V.

I would like to thank the technical staff of the various laboratories I have been in contact with during the last five years: Shirley Shechosky for introducing me to the wonders of protein purification and FPLC, Mike Nattriss and Mike Carpenter for carrying out the amino acid analyzer and the protein sequencer runs, Roger Bradley

for the invaluable assistance in the preparation of the carbon replicas for the electron microscopy and the photographs presented in this thesis, Perry d'Obrenan for tutoring me on the use of the densitometer, Leslie Hicks for carrying out the analytical centrifugation runs and Kim Oikawa for preparing the samples and performing the circular dichroism spectra.

Special thanks are due to the chairman of Biochemistry Dr. Bill Bridger for his encouragement and his sincere interest in my scientific progress and in my post-graduate programs.

I was fortunate enough to have the chance to interact and become an adopted member of the Microbiology Department. I wish to thank them for allowing me to be part of their academic as well as recreational events, such as the Christmas party and the soccer "Micronet" team. In particular, I would like to thank Dr. Susan Jensen for the use of her FPLC system; without her generous offer much of the work presented in this thesis would have not been accomplished. I would also like to thank Joan Hoppenbrouwers for her help on many occasions, too many to list.

I have developed many lifelong friendships during the past few years and I would like to thank the friends I have made; Alberto and Silvia Severini, Amandip K. Utal, Charitini Orphanidou, Francoise Arnaud, Shurjo Ghosh, Ekaterini Nicolarea, Anita Sielecki, Alessandro Butti, without your continuous support and faith in my abilities I would have not succeded. In particular, my most sincere gratitude goes to Dr. Sophia H. Zhou for looking after me for the last few months like the older sister I never had.

The support from the Alberta Heritage Foundation for Medical Research in the form of a studentship and a research allowance is gratefully acknowledged.

TABLE OF CONTENTS

Page

Abstract i

Acknowledgments iii

Table of Contents vi

List of Tables xii

List of Figures xiii

List of Abbreviations xvii

Chapter I. Introduction 1

- A. Bacterial Conjugation 1
- B. IncF Plasmids 4
  - 1. Recognition and attachment to the recipient cell 7
  - 2. Mating pair stabilization 8
  - 3. Conjugal DNA metabolism 9
  - 4. Surface exclusion 15
  - 5. Regulation of expression 16
  - 6. Conjugal DNA replication 17
- C. Research Objectives 18
- D. Bibliography 21

Chapter II. Characterization of the *oriT* Region of the IncFV Plasmid pED208 33

- A. Introduction 33

Page

B. Experimental Procedures 35

1. Bacterial strains and plasmids 35

2. Media 35

3. Transfer efficiency assay 35

4. Protein over-expression and labeling 36

5. Cell fractionation and membrane preparation 37

6. Crude cell extract preparation 38

7. TraM protein purification 38

8. Antisera preparation 39

9. Transfer of protein from SDS-polyacrylamide gels to nitrocellulose 40

10. Immunological detection of proteins on nitrocellulose paper 40

11. Protein composition and sequence 41

12. Recombinant DNA techniques 42

13. Preparation of labelled fragments 42

14. DNA-protein complex electrophoretic mobility retardation assay 42

15. DNase I footprinting 43

C. Results 43

1. Sequence analysis of the *oriT* to *traY* region of pED208 43

2. Determination of the minimal sequence required for transfer 49

3. Overproduction and analysis of the TraM and TraJ proteins 51

Page

4. Immunological characterization of TraM and TraJ proteins 53
5. TraM protein purification 53
6. DNA electrophoretic mobility retardation assays 55
7. DNA footprinting of the *oriT* region with the TraM protein 55
D. Discussion 65
E. Bibliography 69

Chapter III. The Protein TraM of the Conjugative Plasmid pED208 Recognizes a Sequence with Dyad Symmetry 76
A. Introduction 76
B. Experimental Procedures 77
1. Oligonucleotide synthesis 77
2. Preparation of labeled oligonucleotides 77
3. Preparation of double-stranded oligonucleotides 78
4. Electrophoretic mobility retardation assays 79
5. Circular dichroism 79
C. Results 79
1. Choice of synthetic oligonucleotides 79
2. DNA electrophoretic mobility retardation assays 84
3. Competition assays 87
4. Circular dichroism 89
D. Discussion 99
E. Bibliography 103

Page

Chapter IV. The TraM Protein of the Conjugative Plasmid F Binds to the Origin of Transfer of the F and ColE1 Plasmids 106

A. Introduction 106

B. Experimental Procedures 107

1. Bacterial strains, media and plasmids 107

2. Cloning and over-expression of the *traM* gene 107

3. TraM purification 110

4. Antisera preparation 110

5. Detection of the end point of the anti-TraM antisera 111

6. Characterization and localization of TraM within cells 112

7. Recombinant DNA techniques 112

8. Characterization of protein-DNA interactions 112

9. Formation of double-stranded synthetic oligonucleotides 113

C. Results 113

1. TraM over-expression and purification 113

2. Immunological characterization of TraM 114

3. TraM-*oriT* complex formation 117

4. DNase I footprinting analysis of the F and ColE1 *oriT* regions with TraM protein 120

5. TraM binding to synthetic oligonucleotides 123

D. Discussion 126

E. Bibliography 136

Page

Chapter V. Physico-chemical Characterization of the Signal Protein TraM from the Conjugative Plasmids F and pED208 141

A. Introduction 141

B. Experimental Procedures 141

1. Prediction of secondary structure using computer programs 141

2. Analytical ultracentrifugation 142

3. Circular dichroism 144

C. Results 147

1. Sequence comparison of the TraM proteins from the conjugative plasmids F and pED208 147

2. Prediction of secondary structure 150

3. Hydrodynamic characterization of TraM 155

4. Spectrophotometric characterization of the TraM proteins 160

5. Immunodetection 163

D. Discussion 163

E. Bibliography 168

Chapter VI. The Interaction of Integration Host Factor with the Origin of Transfer of the IncFV Plasmid pED208 172

A. Introduction 172

B. Experimental Procedures 173

1. Plasmids 173

Page

2. Recombinant DNA techniques 174

3. Fragment labeling 174

4. Electrophoretic mobility retardation assay 174

5. Calculation of the binding constants 175

6. DNase I protection analysis 176

7. Electron microscopy 176

C. Results 177

1. Electrophoretic mobility retardation assays of the *oriT* region incubated with IHF 177

2. DNase I protection analysis of the *oriT* region incubated with IHF 180

3. Electron microscopy characterization of the IHF-*oriT* complexes 185

4. The interactions of IHF and TraM with the *oriT* region are not co-operative 188

D. Discussion 190

E. Bibliography 198

Chapter VII. Final Conclusions 204

A. pED208 Plasmid 204

B. F Plasmid 209

C. The Icarus Model 214

D. Bibliography 220

LIST OF TABLES

Table Page

II. 1. Chimeric plasmids used in this chapter and their ability to be mobilized by pBF101 50

III. 1. Synthetic oligonucleotide sequences 83

III. 2. Summary of the interactions of TraM with the different oligonucleotides 92

III. 3. Ellipticity values of DNA-protein complexes at 220 nm 98

IV. 1. Bacterial strains 108

V. 1. Amino acid sequence homology among the five prototypes of TraM 148

LIST OF FIGURES

Figure Page

I. 1. Genetic map of the transfer region of the conjugative plasmid F 5

I. 2. Diagram of bacterial conjugation 6

II. 1. Sequence of the first 1782 bp of pED208 after the *Eco*RI site 46

II. 2. Autoradiography of an SDS gel of cells labeled with ^{35}S-methionine expressing TraM and TraJ proteins 52

II. 3. Localization of TraM protein in *E. coli* JC3272/pED208 54

II. 4. SDS-polyacrylamide gel (15%) electrophoresis of samples taken from the steps in the purification of TraM and stained with Coomassie blue 56

II. 5. Binding of TraM protein to the *oriT* region 57

II. 6. DNase I protection of the *oriT* region in the presence of TraM 60

II. 7. DNase I protection of the *oriT* region in the presence of purified TraM 63

II. 8. Summary of the interactions between TraM protein and the 151-360 bp region of *oriT* of pED208 64

III. 1. Linear map of the *oriT* region of the conjugative plasmid pED208 and nucleotide sequence of the TraM binding sites 82

III. 2. Binding of TraM to the single-stranded oligonucleotide **T-T** 85

III. 3. Binding of TraM to the oligonucleotide pairs **T-A** duplex, **T-T/A-A** and **T-T-T/A-A-A** 86

III. 4. Competition of TraM binding to the double-stranded oligonucleotides 88

Figure Page

III. 5. Inhibition of TraM binding to the *oriT* region by the synthetic oligonucleotides 91

III. 6. CD spectra of the interaction of TraM with the single-stranded oligonucleotide **T-T** 95

III. 7. CD spectra of the interaction of TraM with the double-stranded oligonucleotide **T-T/A-A** 97

III. 8. Model of TraM binding to the *oriT* region of pED208 102

IV. 1. Linear map of the clones used in this Chapter 109

IV. 2. Purification of the F TraM protein 115

IV. 3. Localization of the TraM protein 116

IV. 4. DNA electrophoretic mobility retardation assay of TraM-DNA complexes 119

IV. 5. DNase I protection analysis of the *oriT* region in the presence of purified TraM 122

IV. 6. DNA electrophoretic mobility retardation assay of the ColE1 *oriT* region synthetic oligonucleotides 125

IV. 7. Summary of the TraM binding sites in the F *oriT* region 130

IV. 8. Summary of the interactions at the F *oriT* region 133

V. 1. Amino acid sequence alignment of the TraM proteins from the conjugative plasmids F and pED208 149

V. 2. Prediction of the secondary structure of the F TraM protein using the Garnier method and surface probability plots 152

V. 3. Prediction of the secondary structure of the pED208 TraM

Figure Page

protein using the Garnier method and surface probability plots 154

V. 4. Silver staining of the sucrose gradients fractions separated by 15% SDS polyacrylamide gel electrophoresis 156

V. 5. Determination of the weight average molecular weight of pED208 TraM 158

V. 6. Determination of the weight average molecular weight of F TraM 159

V. 7. Ultra-violet scanning of the purified TraM proteins 161

V. 8. Far ultra-violet circular dichroism spectra of TraM proteins 162

V. 9. Near ultra-violet circular dichroism spectra of F and pED208 TraM 164

V. 10. Immunoblot of the purified F and pED208 TraM proteins incubated with anti-F TraM antisera 165

VI. 1. Linear map of the fragments used in this Chapter 179

VI. 2. Binding of IHF to the *oriT* region of pED208 181

VI. 3. DNase I protection of pED208 *oriT* region incubated with IHF 183

VI. 4. Electron microscopy of IHF-*oriT* region complexes 187

VI. 5. Position of the bends in the 742 bp and 648 bp fragments as observed by electron microscopy 189

VI. 6. DNase I protection of pED208 *oriT* region incubated with IHF and TraM 192

VI. 7. Diagram of the interactions at the *oriT* region of the IncF plasmids pED208, F and R100 194

VII. 1. Amino acid sequence alignment of the five TraM prototypes 207

VII. 2. The Icarus model 215

LIST OF ABBREVIATIONS

A	adenine
ATP	adenosine triphosphate
bp	base pair
BSA	bovine serum albumin
C	cytosine
°C	centigrade degree
CD	circular dichroism
cm	centimeter (10^{-2} meter)
cpm	radioactive counts per minute
Col	colicin producing plasmid
Δ	deletion
dATP	deoxyadenine triphosphate
dCTP	deoxycytidine triphosphate
deg	degree
DNA	deoxyribonucleic acid
DNase I	deoxyribonuclease I
DTT	dithiothreitol
ε	molar extinction coefficient
EDTA	ethylenediamine tetraacetic acid
EM	electron microscopy
F	F conjugative plasmid
fin	fertility inhibition
fmol	femtomole (10^{-15} mole)
g	acceleration of gravity

G	guanine
HEPES	N-(2-Hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
Inc	incompatibility
IPTG	isopropyl-β-thiogalactopyranoside
K_b	equilibrium binding constant
kb	kilobase (10^3 base)
kDa	kiloDalton (10^3 Dalton)
ln	natural logarithm
ml	milliliter (10^{-3} liter)
mA	milliAmpere (10^{-3} Ampere)
min	minute
μl	microliter (10^{-6} liter)
nm	nanometer (10^{-9} meter)
OD	optical density
PAG, PAGE	polyacrylamide gel, polyacrylamide gel electrophoresis
pH	negative logarithm of the hydrogen ion activity
pI	isoelectric pH
pmol	picomole (10^{-12} mole)
PMSF	phenylmethylsulfonylfluoride
θ_λ	ellipticity at the wavelength λ
R	resistance conjugative plasmid
rpm	revolution per minute
SDS	sodium dodecyl sulfate
St	molecular weight standard
T	thymine

TFE	trifluoroethanol
Tris-HCl	Tris(hydroxymethyl)aminomethane hydrochloride
UV	ultra-violet
%	percent
≈	approximately

In the text of the thesis an italicised name indicates a gene locus, whereas the normal print gene locus name refers to the gene product, i.e. TraM is the gene product of *traM*.

In the text of the thesis, other than strain genotypes, a mutant allele of a gene is indicated with an upper right $^-$, i.e, *lac*$^-$.

In the text of the thesis the terms *oriT*, *oriT* site and nick site indicate the phosphodiester bond that is cleaved by the site- and strand-specific endonuclease, whereas the terms *oriT* region and OriT indicate the DNA sequence containing *oriT* and the sites of interaction of the factors involved in DNA transfer.

In the text of the thesis the term upper strand indicates the strand of the conjugative plasmid which is retained in the donor cell, whereas lower strand indicates the transferred strand.

CHAPTER I

Introduction

Conjugation involves the transfer of a single-stranded DNA molecule between bacterial cells in close contact with each other. The major implication of such a process is that a wide pool of information is accessible to a variety of organisms. Information in the form of metabolic pathways of simple carbohydrates to complex hydrocarbons, from resistance to antibiotics and drugs to toxin production is shared. Thus conjugation contributes to the genetic variability and adaptability of bacteria to adverse environmental conditions (Mazodier and Davies, 1991). Conjugation requires the following functions: production of a sensor able to detect a recipient cell and establish the mating contact, stabilization of the contacts between the mating pair, activation of DNA transfer, and regulation of the expression of the conjugative machinery. All of the above-mentioned functions are encoded by the conjugative plasmids, which also possess the site where DNA transfer originates, called the *oriT*.

A. Bacterial Conjugation

Bacterial conjugation was first observed by Joshua Lederberg in 1946. To explore the possibility of sexual recombination in bacteria, he tried to complement Edward Tatum's *Escherichia coli* double mutants using the *E. coli* wild type strain K12 (Lederberg and Tatum, 1946). The factor responsible for the complementation was called the fertility (F) factor (Hayes, 1953). For a long time F was simply considered a point in the *E. coli* chromosome from which the transfer started. It was later discovered that F is a circular DNA molecule of 100 kb capable of integration into the bacterial chromosome (Marmur *et al.*, 1961, Freifelder, 1968).

Four highly unlikely circumstances allowed the first observation of conjugation. First, the *E. coli* K12 strain carried the conjugative plasmid F; this was rare in the pre-antibiotic era. Second, the F factor was derepressed as a result of the insertion of an insertion element (IS3) in the F genome, consequently, "fertility inhibition control" was inactivated (Willetts and Skurray, 1980); on average every of such F^+ cells is conjugation proficient (Brinton *et al.*, 1964). Third, F possesses a collection of insertion sequences that allow site-specific recombination with the *E. coli* chromosome, thus creating the Hfr phenotype necessary for chromosome transfer. Fourth, the double mutant strain had lost its F plasmid and could act as a recipient for F.

The discovery of bacterial conjugation was a cornerstone in the study of *E. coli* genetics and contributed to the understanding of the circular structure (Jacob and Wollman, 1961) as well as genetic mapping of the bacterial chromosome.

The F factor was the first of a large number of conjugative systems to be identified, although the detailed mechanisms involved in each system vary significantly. Conjugative systems include the resistance factors or R plasmids which are responsible for the spread of resistance to antibiotics among pathogenic bacteria (Taylor, 1989), the tumor inducing (Ti) plasmids of *Agrobacterium tumefaciens* which infect plant cells and cause crown gall formation (Zambryski, 1989), and the bacteriocin-producing plasmids, some of which are not capable of autonomous transfer but can be mobilized by other conjugative factors.

There are two major criteria for the classification of conjugative plasmids. The first is based on their ability to co-exist in the bacterial cell. If Incompatibility (Inc) is defined as the result of controls occurring at the level of replication and partitioning of the plasmid (Austin and Nordström, 1990), then two plasmids that cannot exist in the same cell belong to the same Inc group. If they can be stably maintained, they are members of different Inc groups (Couturier *et al.*, 1988, Novick, 1987). Members of the same Inc group usually share regions of DNA homology that correspond to related conjugation systems. More than 20 Inc groups have been identified and their members have been isolated from a wide range of Gram-negative bacteria (Taylor, 1989) and from five Gram-positive genera (Mazodier and Davies, 1991). Among the Inc groups most extensively studied are the members of the IncF and IncP groups. Members of the first group are called IncF because of their similarities to the F plasmid with respect to their transfer regions (Willetts and Maule, 1986). Five IncF groups have been identified, the prototypes of each group are: F, ColB4, R1, R100 and F_0*lac* (Willetts and Maule, 1986, Finlay *et al.*, 1986a, 1986b). Members of the incP group are characterized by their ability to transfer and be stably maintained in a wide range of hosts, *i. e.* they are promiscuous.

The second criterion for classifying conjugative plasmids is the host range. Conjugative plasmids capable of transfer within the same or closely related species are called narrow host range, whereas plasmids able to propagate among different species and genera are called broad host range or promiscuous plasmids. The best characterized narrow host range plasmids are the IncF plasmids, and among the promiscuous plasmids the IncP plasmids are the best known. The utility of this criterion for plasmid classification has been reduced by the discovery of F-mediated DNA transfer between *E. coli* and the Gram-negative bacterium *Pseudomonas aeruginosa* (Guiney, 1982), and *E. coli* and the yeast *Saccharomyces cerevisiae* (Heinemann and Sprague, 1989). These studies show that intergeneric and interkingdom transfer is possible if the mobilized plasmids carry a vegetative origin of replication recognized by the recipient host. This implies that host range selectivity is the result of plasmid replication and maintenance functions rather than of recognition of a specific receptor on the surface of the recipient.

B. IncF Plasmids

In the F system the genes involved in conjugation, including *oriT*, are clustered in a single region called the transfer (*tra*) region (Ippen-Ihler and Minkley, 1986). Early genetic investigations coupled with the recently completed sequence of the 34 kb *tra* region of F have identified more than thirty genes and open reading frames (Fig. I. 1). The region containing *oriT* and the site of interaction of the *tra* gene products involved in DNA transfer is called the *oriT* region. It is located at one end of the transfer region on a 1.1 kb *Bgl*II fragment (66.7 kb in the F map) such that during the conjugal DNA transfer, the *tra* region is the last to leave the donor cell (Willetts, 1972; Thompson and Achtman, 1978; Achtman *et al.*, 1971). Despite the wealth of genetic information accumulated, the detailed mechanism of conjugal DNA transfer is still unclear. Figure I. 2 shows a schematic representation of conjugation and the

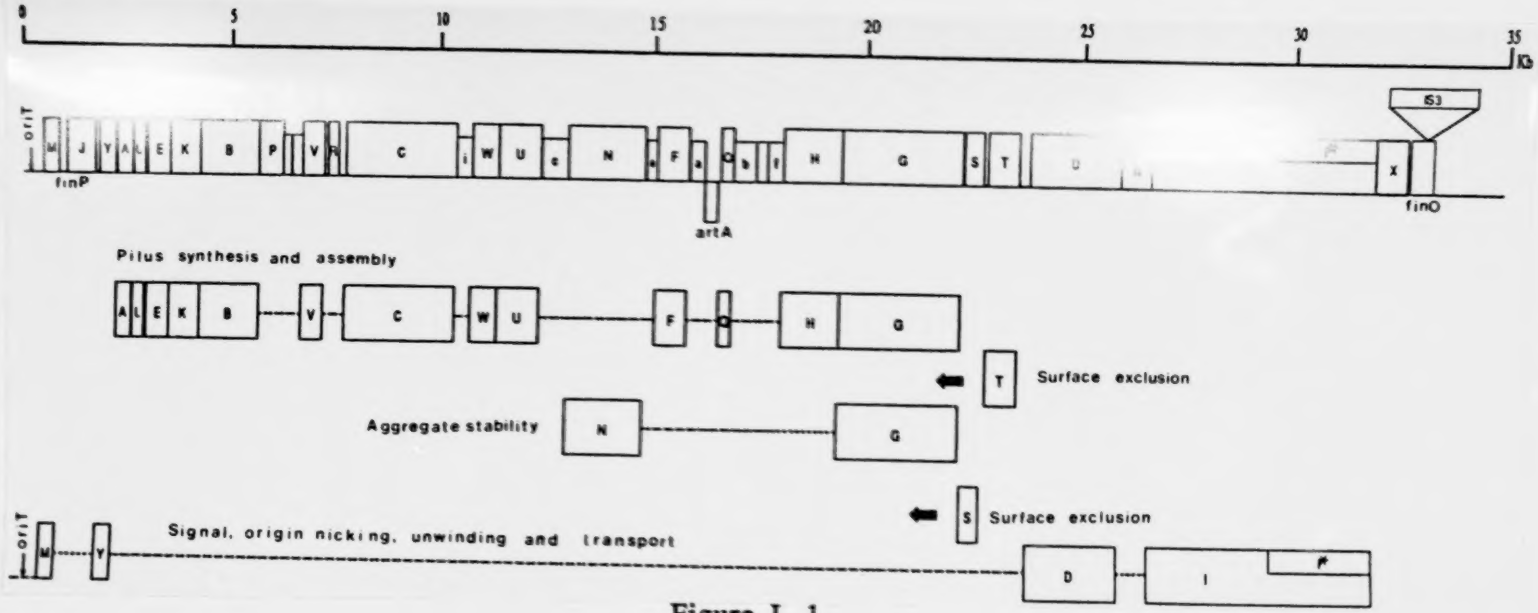

0
5
10
15
20
25
30
35
Kb
oriT
M
J
Y
A
L
E
K
B
P
V
R
C
i
W
U
e
N
F
a
Q
b
I
H
G
S
T
D
X
IS3
finP
artA
finO
Pilus synthesis and assembly
Surface exclusion
Aggregate stability
Surface exclusion
Signal, origin nicking, unwinding and transport

Figure I. 1

Genetic map of the transfer region of the conjugative plasmid F. The genetic map of the transfer region is shown with the genes and open reading frames identified. Transcription of the operon proceeds rightward from the *traY* promoter except for *artA* and *finP*. *traM* and *traJ* have promoters that direct transcription in the same direction as p_Y. The genes whose functions are known are grouped accordingly into: pilus synthesis and assembly, surface exclusion, aggregate stability and signal, origin of nicking, unwinding and DNA transport. This diagram is from Dr. Karin Ippen-Ihler, Texas A&M University. Note that the *traX* gene function was not known at the time this map was released.

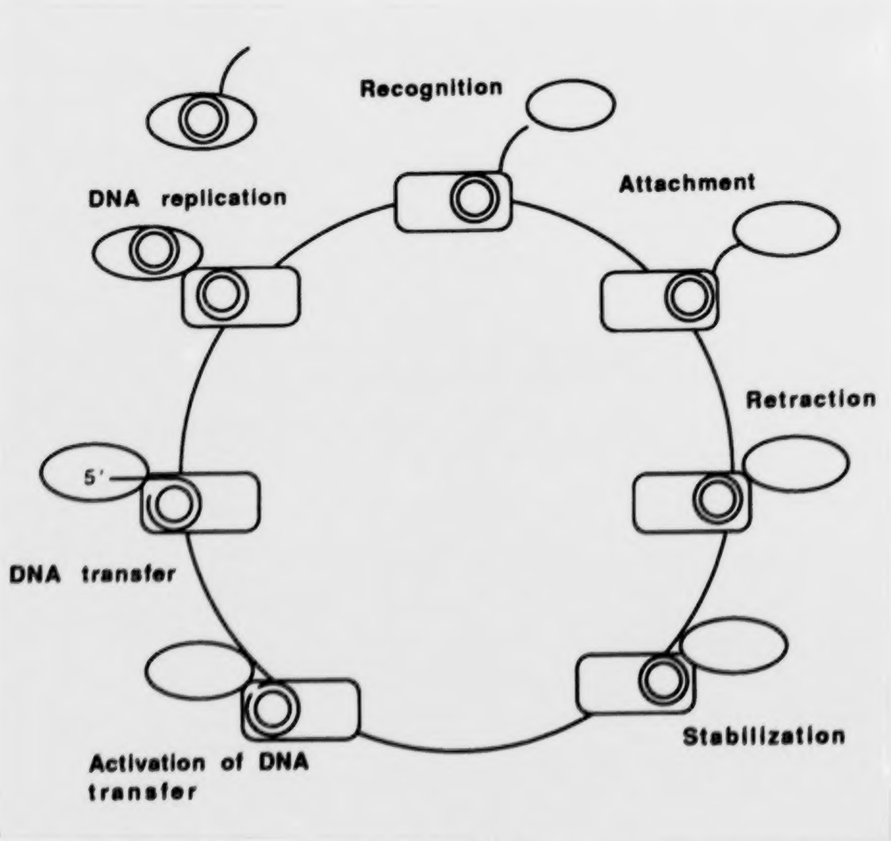

Recognition
Attachment
Retraction
Stabilization
Activation of DNA transfer
DNA transfer
5'
DNA replication

Figure I. 2

Diagram of bacterial conjugation. The process is presented according to the current model and each step is discussed in the text.

following sub-sections will introduce each step of the process and the genes involved.

1. Recognition of and attachment to the recipient cell

In F and F-like plasmids the sensor is a hair-like appendage, called the sex pilus. In F-mediated conjugation the receptor of the pilus is thought to be the outer membrane protein A (OmpA) (Willetts and Maule, 1986). The pilus is a homopolymer of the 7 kDa pilin subunit (Ippen-Ihler *et al.*, 1984) encoded by the *traA* gene as pre-pro-pilin of 13-14 kDa (Frost *et al.*, 1984; Minkley *et al.*, 1976). The mature pilin subunits are arranged with a fivefold rotational axis of symmetry around the pilus axis (Folkhard *et al.*, 1979) and the interactions among subunits are likely to occur via hydrophobic bonds, since pili can be dissociated into sub-units by treatment with 0.1% SDS (Achtman *et al.*, 1978b).

In contrast to the simplicity of the pilus biochemical composition, more than 14 *tra* gene products are involved in the synthesis and assembly of the pilus (Fig. I. 1). The functions of two of these gene products have been elucidated: the pre-pro-pilin sub-unit undergoes amino-terminal cleavage of the first 51 residues in one or more steps mediated by the *traQ* gene product, the amino-terminus of the mature pilin is then acetylated by the *traX* gene product (Ippen-Ihler and Maneewannakul, 1991).

The pilus is also the receptor for two groups of phages: (1) F-specific RNA phages that bind to its lateral surface of the pilus (Brinton *et al.*, 1964), and (2) filamentous single-stranded DNA phages that attach to the tip (Caro and Schnös, 1966). Sensitivity to these phages has been a powerful tool in the genetic characterization of the *tra* mutants.

The pilus is essential for the establishment of mating pairs, since sodium dodecyl sulfate (SDS) treatment of cells before mating pair stabilization totally blocks transfer (Willetts and Skurray, 1987; Panicker and Minkley, 1985). The role the pilus plays for conjugal DNA transfer is, however, controversial. The nature of the pore through which DNA moves into the recipient cell has not been elucidated, and conjugation between cells physically separated by porous membranes has been demonstrated, suggesting that transfer of DNA may occur through the pilus (Harrington and Rogers, 1990).

Upon attachment the pilus retracts, bringing the cells into wall-to-wall contact to form the mating pair. Very little is known about the mechanism of retraction or the molecular nature of the region of junction between mating cells. Some data suggest that retraction may occur by depolymerization of the pilus into the donor cell where the pilin subunits change from an axial arrangement to a sheet distribution creating the junction between mating cells (Folkhard *et al.*, 1979).

2. Mating pair stabilization

The next step in conjugation is the stabilization of the mating pair which depends on the action of the *traN* gene and the 3' end of the *traG* gene (Manning *et al.*, 1981). The biochemical mechanism of stabilization is not known, but there are genetic data suggesting that OmpA of the recipient cell may be involved (Willetts and Skurray, 1980). Once the mating pair is stable (SDS- and shear-resistant), a signal is transmitted to the DNA to start transfer. Everett and Willetts (1980) used λ*oriT*⁺ transducing phages to map the nick site *in vivo*. The rationale for this approach was as follows: if F⁺ cells were infected with lambda phages carrying the *oriT* site, then the phage progeny would be nicked at *oriT*. They found that the products of *traY* and *traZ* genes are required for nicking, whereas previous genetic analyses had identified the *traM*, *traI*, *traD* genes and the *oriT* region as well as *traY* and *traZ* as essential functions for the conjugal DNA transfer (Kingsman and Willetts, 1978).

3. Conjugal DNA metabolism

Upon formation of a stable mating pair, the putative signal protein TraM is thought to trigger the transfer process. The gene products of *traI* (DNA helicase I) (Abdel-Monem *et al.*, 1983) and *traD* unwind and "pump" the DNA through the conjugal bridge in a 5' to 3' direction (as reviewed by Willetts and Skurray, 1987). Although the 3' end of the transferred DNA is a free hydroxyl group (Everett and Willetts, 1980), it can not be ligated to its 5' end. This led to the suggestion that the latter was either partially hydrolyzed or modified (Thompson *et al.*, 1989). Therefore Willetts and Wilkins proposed a model for conjugal transfer in which the 5' end of the transferred strand is covalently bound to the TraYZ "nickase" and the protein-DNA complex is anchored at the junction between the mating cells by TraD. As a result, the single-stranded DNA is transferred as a loop into the recipient cell, where a complementary strand is synthesized. At the end of the transfer process the 3' end was proposed to be ligated to the 5' end before leaving the donor cell (Willetts and Wilkins, 1984).

The sequences of several *oriT* regions reflect the differences in the *tra* genes involved in DNA metabolism. Complementation analysis identified four *oriT* classes (Willetts and Maule, 1985; Finlay *et al.*, 1986a) that show similar sequence organization, such as a conserved region rich in GT (in the nicked strand), a semiconserved region of high AT content, and a variable region containing several direct and inverted repeats. *In vivo* experiments mapped three potential *oriT* sites in the GT-rich region of F. Eventually, this region was proposed to be the site of the nick in the other three *oriT* classes (Finlay *et al.*, 1986a). More recently, the *oriT* site was mapped *in vivo* at position 137 from the *Bgl*II site at 66.7 kb in the F map (Thompson *et al.*, 1989), and at position 138 *in vitro* (Reygers *et al.*, 1991); the discrepancy was attributed to the stereochemical hindrance of the last nucleotide to the sequencing reaction, as a result of the modification of the 5' end (Reygers *et al.*, 1991). The AT-rich region and the series of direct and inverted repeats were postulated to be the sites of binding of TraY and TraM proteins (Finlay *et al.*, 1986a).

Recent biochemical experiments have been directed towards understanding the molecular and biochemical mechanism of initiation of DNA transfer. In 1988, Traxler and Minkley mapped the *traZ* locus within the *traI* gene. They later found that the F *traZ* locus is associated with the 5' end of the *traI* gene (Traxler and Minkley, 1987, 1988); therefore, they postulated a bifunctional role for the TraI protein as a DNA helicase I and as a site-specific DNA "nickase". TraI catalyzes a nucleoside 5'-triphosphate-dependent unwinding reaction of partial duplex DNA (Lahue and Matson, 1988), but DNA helicase I is unable to unwind completely duplex DNA or nicked molecules (Abdel-Monem *et al.*, 1977). It requires a single-stranded DNA region approximately 200 bp long (Kuhn *et al.*, 1979).

Two laboratories have demonstrated the bifunctional activity of F TraI as the site- and strand-specific endonuclease that nicks *oriT* as well as a helicase (Reygers *et al.*, 1991; Matson and Morton, 1991). Nicking at *oriT* occurs when the plasmid contains negative supercoils and in the presence of Mg^{2+}. The *in vitro* reaction does not require TraY (Matson and Morton, 1991) in contrast with the prediction of Willetts and Wilkins model (1984) and the finding that *traY* and *traZ* are necessary for nicking λ*oriT* phage genome; however, the efficiency of the *in vitro* reaction is between 50% and 70% of its efficiency *in vivo* (Everett and Willetts, 1980). Electron microscopy revealed aggregates of TraI attached to *oriT*-containing fragments, thus identifying the source of the modification of the 5' end of the nicked strand (Reygers *et al.*, 1991). Electrophoretic mobility retardation assay and DNase I protection analyses indicate the F TraY protein interacts with the DNA at a site 60 bp downstream of the *oriT* site. However TraY does not nick the DNA in these experiments (Lahue and Matson, 1990). TraI fails to protect the *oriT* region from DNaseI digestion (Matson and Morton, 1991).

In the IncFIV plasmid R100, the site- and strand-specific endonuclease activity of the TraY-TraI complex was demonstrated to occur in reaction conditions similar to the F system (Inamoto *et al.*, 1991). Furthermore, the *oriT* site of R100 was mapped, and corresponds to the site predicted by sequence comparison (Finlay *et al.*, 1986a).

The TraM proteins of the F-like plasmids R1 and R100 have also been characterized (Schwab *et al.*, 1991; Abo *et al.*, 1991). Both proteins interact with their homologous *oriT* regions at sites predicted to be the TraM binding sites (Finlay *et al.*, 1986a). In the R1 system, two large regions of interaction separated by 12 bases were identified. These regions appear to have multiple binding sites for TraM (Schwab *et al.*, 1991) whereas, in the R100 *oriT* region, four binding sites very close to each other were characterized (Abo *et al.*, 1991). In both cases the binding sites are located approximately 100 bp away from the *oriT* sites and overlap the *traM* gene promoter regions, suggesting autoregulation of *traM* expression. However, TraM function in DNA transfer remains elusive. Abo *et al.* found the over-expressed TraM associated with the membrane and dispensable from the *in vitro* nicking reaction. They proposed that TraM binds to the *oriT* region of R100 and anchors the DNA to the membrane (Abo *et al.*, 1991), whereas Schwab *et al.* suggested an interaction between TraM and TraY proteins bound to adjacent sites in the *oriT* region of the R1 plasmid (Schwab *et al.*, 1991).

The interaction of Integration Host Factor (IHF) protein with the *oriT* region was analysed in the R100 and F systems. These studies were triggered by earlier reports in which IHF$^-$ strains showed between 100- and 450-fold decreased transfer efficiency (Dempsey, 1987; Gamas *et al.*, 1987). IHF is a histone-like protein of *E. coli* that plays an essential role in site-specific recombination of phage λ into the bacterial chromosome (Friedman, 1988; Drlica and Rouviere-Yaniv, 1987). IHF, binds to specific sequences in AT-rich regions, and bends the DNA by more than 140° (Yang and Nash, 1989), bringing the binding sites of the Integrase protein into close proximity. Integrase catalyzes recombination (Ross *et al*, 1979).

In vitro studies identified two sites of IHF interaction in the *oriT* region. The first site of high affinity (IHFA) is between the *oriT* site and the TraY binding site; the second site is located in the variable region towards the *traM* promoter (Tsai *et al.*, 1990; Dempsey and Fee, 1990). These results suggest that a multiprotein complex acts at *oriT*, and that IHF binding to the high affinity site would bring the site of interaction of TraY into close proximity with the *oriT* site. The second IHF binding site was proposed to have an effect on *traM* expression (Inamoto *et al.*, 1990).

IHF has been shown to have an effect on the expression of some *tra* genes. In the derepressed conjugative plasmid R100-1 two transcriptional products of the *traM* gene were identified: a 705 nucleotide transcript and a shorter transcript 562 nucleotides long. The latter is initiated 148 bp downstream of the start site of the longer transcript, and it is not under the control of any R100-1 gene products. Both transcripts terminate inside the *traJ* open reading frame, although the exact position of the 3' end is not known (Dempsey, 1989). Synthesis of the longer transcript is increased by the presence of the *traJ* gene, whose gene product is believed to activate expression of the *traY* operon (see section **5. Regulation of expression**). However, in these experiments *traM* expression does not reach the levels observed in the native conjugative plasmid, suggesting that regulation of *traM* expression is dependent on *traJ* and some other unknown factors. The product of the longer transcript encodes a polypeptide of relative mobility of 10,000 daltons on SDS-PAGE, and corresponds to TraM. No translational product of the shorter transcript was identified in maxicells. It was proposed that the shorter transcript (*finM*) interacts with the *finP* product, an anti-sense RNA molecule complementary to the 5' untranslated region of the *traJ* gene (see section **5. Regulation of expression**) and prevents the establishment of the FinOP system which represses the *traJ* gene (Dempsey, 1989).

Evidence for the effect of IHF on the expression of *traM* was presented by Dempsey and Fee (1990) in the R100-1 system. They showed that full-length transcripts of the *traM* gene are reduced in *E. coli* IHF$^-$ mutants carrying the cloned R100-1 *traM* gene, whereas shorter transcripts of mobility corresponding to *finM* accumulate in IHF$^-$ strains. The reasons for the transition remain unclear (Dempsey and Fee, 1990). Furthermore, *traM* transcripts appear even in the absence of the *traJ* gene, indicating that *traM* expression is independent of *traJ*.

Transcription of at least one other gene in the R100 *oriT* region was shown to be dependent on IHF (Dempsey and Fee, 1990). The *oriT* region of IncF plasmids is located between the *traM* and the gene *X* promoters. The latter is on the nicked strand 3' from the nick site. Thus it is the first gene to enter the recipient cell and be transcribed. The function of gene *X* is not known. It contains an open reading frame that could encode a basic protein of 19.4 kDa (Fee and Dempsey, 1990). Probes specific for gene *X* identified two species of RNA that were shorter than full-length transcripts in IHF$^-$ mutants. It was proposed that the two short transcripts contain secondary structures which prevent their immediate degradation, thus allowing their detection on Northern blots. The size and amount of gene *X* transcript in wild type IHF$^+$ cells is not known (Dempsey and Fee, 1990).

The multiprotein complex model implies that the distance between the functional domains of the *oriT* region might be important. This was demonstrated by the detailed dissection of the *oriT* region of F performed by Fu *et al.* (1991). Deletion mutants of the *oriT* region were tested for their ability to be nicked and promote transfer. Two sequences, essential for the nicking reaction, were mapped in the second half of the conserved region extending into the AT-rich region. They corresponded to the *oriT* site and IHFA. Two more sequences were identified in the region containing the inverted and direct repeats that are essential for transfer. Removal of these sequences reduced transfer efficiency by four orders of magnitude. This regions had been proposed to contain the sites of interaction of the *traM* gene product (Finlay *et al.*, 1986a). As expected, the distance between the functional domains is important for efficient nicking and DNA transfer. Insertion of linker sequences between the functional determinants identifies a value of 12.25 bp for the optimum inter-domain distances. This value is higher than the common range for DNA turns (10.5-11 bp), suggesting the possible underwinding of the DNA upon protein binding (Fu *et al.*, 1991).

4. Surface exclusion

Conjugation among donor cells is prevented by the surface exclusion function (Lederberg and Tatum, 1946). Two *tra* genes are responsible for this phenomenon, *traS* and *traT* (Achtman *et al.*, 1980). F TraS is an 18 kDa cytoplasmic membrane protein, whereas TraT is a 25 kDa lipoprotein (Perumal and Minkley, 1984) located in the outer membrane (Achtman *et al.*, 1977, 1978a). TraS function is not understood but it may involve the blocking of DNA transfer (Manning and Achtman, 1979). TraT interacts with OmpA, the receptor of the pilus (Achtman *et al.*, 1977), thus preventing recognition and attachment (Riede and Eschbach, 1987).

5. Regulation of expression

The regulation of expression of most of the *tra* genes is achieved by the *traJ-finOP* system. Most of the *tra* genes are organized in the *traY* operon (Fig. I. 1) under the control of the p_Y promoter. This promoter has no homology with the consensus promoter sequence in prokaryotic organisms, implying the presence of a control mechanism (Finnegan and Willetts, 1973). Genetic studies suggest TraJ as a positive regulator of the expression of the p_Y promoter (Willetts, 1977; Gaffney *et al.*, 1983).

To test the dependence of the F *traY* promoter activity on TraJ, Silverman *et al.* (1990) created constructs of the *traY* promoter region fused to the *lacZ* gene. TraJ produces an ≈ 20 fold increase in β-galactosidase activity. The activity of p_Y is also under the control of two host factors, SfrA, and, at a minor level, IHF (Silverman *et al.*, 1991). SfrA is involved in the repression of chromosomal genes under aerobic respiration control, and SfrA$^-$ mutants show an ≈15 fold reduction of the p_Y promoter activity. IHF was shown earlier to affect efficiency of transfer of the F plasmid (Gamas *et al.*, 1987). *In vitro* analyses of the *traY* promoter also show that TraY protein interacts with this region (Lahue and Matson, 1990). These results suggest a complex mechanism of control of the *traY* promoter which is dependent on plasmid as well as host functions. This is not surprising since the *traY* promoter is the major switch for the expression of the conjugative machinery.

traJ gene expression is down regulated by the action of the fertility inhibition system FinOP (Finnegan and Willetts, 1973) which has two components: FinP, an antisense RNA molecule complementary to the untranslated region of the *traJ* messenger RNA; and FinO, the product of the distal transfer gene *finO* (Fig. I. 1). F is naturally derepressed because of the insertion of an IS3 element in its *finO* gene (Fig. I. 1). However the *finO* gene of compatible plasmids can cause repression of F (Finnegan and Willetts, 1973). FinO binding to FinP increases the half life of the RNA molecule from 7.5 to more than 40 min (S. Lee, MSc thesis, University of Alberta) even in the absence of the *traJ* gene. The detailed mechanism of FinOP action on *traJ* expression is controversial and the following model has been proposed: FinO increases FinP half-life such that it binds *traJ* messenger RNA to trigger degradation of the double-stranded RNA by the action of host RNase(s) (van Biesen and Frost in press). Testing of this model awaits the purification and characterization of the FinO protein.

6. Conjugal DNA replication

Both DNA strands are replicated as part of the conjugative event, and the nature of the factors involved in the process has been investigated. Replication of both the donor and transferred strand is dependent on DNA polymerase III as shown by the analysis of *dnaE* temperature sensitive mutants (Kingsman and Willetts, 1978; Wilkins and Hollom, 1974). In the donor cell, replication could use the free 3' hydroxyl group as a primer in accordance with the rolling circle mechanism. Or, an RNA molecule synthesized by RNA polymerase could be used to prime, since in F and F-like plasmids conjugal DNA synthesis was found to be rifampicin sensitive (Kingsman and Willetts, 1978). The synthesis of the strand complementary to the transferred strand was proposed to be discontinuous, with short RNA primer molecules synthesized by *dnaG* gene product (Wilkins and Hollom, 1974).

Unsuccessful attempts were made to locate TraI or the plasmid-encoded single-stranded DNA binding (ssf) proteins (Kolodkin *et al.*, 1983) transferred together with the DNA into the recipient cell. Only a small amount of a 92 kDa protein identified as TraD is detected in the outer membrane fraction of recipient cells (Rees and Wilkins, 1990). These results support the model in which the 5' end of the transferred strand is bound to TraI and anchored to the membrane at the junction between the mating cells during transfer. There is no evidence for a primase gene in F and F-like plasmids, unlike the promiscuous plasmids RP4 (Rees and Wilkins, 1990) and ColI-P9 (Rees and Wilkins, 1989).

It is clear from the overview presented (see sub-section **3. Conjugal DNA metabolism**) that several proteins act at the *oriT* region to initiate and control the DNA transfer process in response to the mating signal. It is believed that the site- and strand-specific TraYI endonuclease continuously nicks and ligates *oriT* like topoisomerase I but without the strand exchange, and that IHF binding between *oriT* and the TraY binding site brings the two sites into close proximity. It has also been postulated that TraM is the signal protein which may interact with two regions in the *oriT* region that are essential for transfer. These are located next to the TraY binding site and may also regulate expression of the *traM* gene. How these proteins interact and "talk" to each other is unclear.

C. Research Objectives

One of the most delicate steps in bacterial conjugation is the initiation of DNA transfer into the recipient cell. Since the transfer process occurs over a period of several minutes it is important that the contacts between the mating cells are stable. Thus, the signalling of the establishment of strong contacts to the multiprotein complex at *oriT* is essential to achieve successful conjugation and, ultimately, efficient propagation of the plasmid.

Among the F-like plasmids, this function was attributed to the *traM* gene. Cells carrying *traM*$^-$ F plasmids (F*lac traM*102) are sensitive to pilus-specific phages but do not transfer DNA (Kingsman and Willetts, 1978). The *traM* locus is found in the region between the *oriT* region and the *traY* promoter, and is therefore the last gene to leave the donor cell during transfer. Nucleotide sequencing of the *traM* gene of F identified an open reading frame, and *in vitro* transcription experiments have mapped the promoter (Thompson and Taylor, 1982). This region is unusual in that it corresponds to the second half of a pair of direct repeats, suggesting modulation of the transcription of this gene. Preliminary biochemical characterization of TraM protein using minicells has localized it in the cytoplasmic membrane fraction (Kennedy *et al.*, 1977) although analysis of the amino acid sequence does not predict the presence of hydrophobic regions (Thompson and Taylor, 1982).

pED208 is a 90 kb conjugative plasmid belonging to the IncFV class, first isolated from a strain of *Salmonella typhi* (Finlay *et al.*, 1983). It is of particular interest because cells that carry the plasmid are multipiliated. Cells carrying the plasmid produce on average 17 pili per cell (Frost *et al.*, 1985). Multipiliation is the result of the insertion of an IS2 element at the beginning of the *traY* gene, thus the *traJ-finOP* control is by-passed (Finlay *et al.*, 1986b). The plasmid does not have extensive DNA homology with F as shown by Southern blotting (Finlay *et al.*, 1984), although the two plasmids encode similar gene products and show a similar arrangement of genes. The nucleotide sequence of the region downstream of the IS2 element identifies the gene for the pilin sub-unit *traA*, three of the genes involved in pilus synthesis, *traL*, *traE*, *traG*, and the genes for surface exclusion *traT* and *traS* (Finaly *et al.*, 1986b; Finlay and Paranchych, 1986). The region upstream the IS2 element is similar to the *oriT*, *traM* and *traJ* regions of F (see Chapter II).

This thesis reports the cloning, over-expression and purification of the *traM* gene product from two IncF plasmids. I have characterized their DNA binding activities and compared the physical properties of the two proteins in order to gain a better understanding of the mechanism of initiation of DNA transfer. In addition, the interaction of Integration Host Factor with the origin of transfer of the pED208 plasmid has been studied.

D. Bibliography

Abdel-Monem, M., Lauppe, H.-F., Kartenbeck, J., Dürwald, H., Hoffman-Berling, H. (1977) Enzymatic unwinding of DNA; mode of action of *Escherichia coli* DNA unwinding enzyme. *J. Mol.Biol.* **110**: 667-685.

Abdel-Monem, M., Taucher-Scholz, G., and Klinkert, M.Q. (1983) Identification of *Escherichia coli* DNA helicase I as the traI gene product of the F sex factor. *Proc. Natl. Acad. Sci. U. S. A.* **80**: 4659-4663.

Abo, T., Inamoto, S., and Ohtsubo, E. (1991) Specific DNA Binding of the TraM Protein to the *oriT* Region of Plasmid R100. *J. Bacteriol.* **173**: 6347-6354, 1991.

Achtman, M., Kennedy, N., and Skurray, R. (1977) Cell-cell interactions in conjugating *Escherichia coli:* role of the TraT protein in surface exclusion. *Proc. Natl. Acad. Sci. U. S. A.* **74**: 5104-5108.

Achtman, M., Manning, P.A., Kusecek, B., Schwuchow, S., Willetts, N. (1980) A genetic analysis of F sex factor cistrons needed for surface exclusion in *Escherichia coli. J. Mol. Biol.* **138**: 779-795.

Achtman, M., Morelli, G., and Schwuchow, S. (1978a) Cell-cell interactions in conjugating *Escherichia coli*: Role of F pili and fate of mating aggregates. *J. Bacteriol.* **135**: 1053-1061.

Achtman, M., Schuwchow, S., Helmuth, R., Morelli, G., Manning, P.A. (1978b) Cell-cell interactions in conjugating *Escherichia coli*: Con- mutants and stabilization of mating aggregates. *Mol. Gen. Genet.* **164**: 171-183.

Achtman, M., Willetts, N.S., and Clark, A.J. (1971) Beginning a genetic analysis of conjugational transfer determined by the F factor in *E. coli* by isolation and characterization of transfer-deficient mutants. *J. Bacteriol.* **106**: 529-5381.

Austin, S., and Nordstrom, K. (1990) Partition-mediated Incompatibility of bacterial plasmids. *Cell* **60**: 351-354.

Brinton, C.C., Gemski, P.,Jr., and Carnahan, J. (1964) A new type of bacterial pilus genetically controlled by the fertility factor of *Escherichia coli* K12 and its role in chromosome transfer. *Proc. Natl. Acad. Sci. U. S. A.* **152**: 776-783.

Caro, L.G., and Schnos, M. (1966) The attachment of male-specific bacteriophage f1 to sensitive strains of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **56**: 776-783.

Couturier, M., Bex, F., Berquist, P.L., Maas, W.K. (1988) Identification and classification of bacterial plasmids. *Microbiol. Rev.* **52**: 375-395.

Dempsey, W.B. (1987) Integration host factor and conjugative transfer of the antibiotic resistance plasmid R100. *J. Bacteriol.* **169**: 4391-4392.

Dempsey, W.B. (1989) Sense and antisense transcripts of *traM*, a conjugal transfer gene of the antibiotic resistance plasmid R100. *Mol. Microbiol.* **3**: 561-570, 1989.

Dempsey, W.B., and Fee, B.E. (1990) Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol. Microbiol.* **4**: 1019-1028.

Drlica, K., and Rouviere-Yaniv, J. (1987) Histonelike proteins of bacteria. *Microbiol. Rev.* **51**: 301-319.

Everett, R., and Willetts, N. (1980) Characterisation of an in vitro system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136:** 129-150.

Finlay, B.B., Frost, L.S., and Paranchych, W. (1984) Localization, cloning, and sequence determination of the conjugative plasmid ColB2 pilin gene. *J. Bacteriol.* **160**: 402-407.

Finlay, B.B., Frost, L.S., and Paranchych, W. (1986a) Origin of transfer of IncF plasmids and nucleotide sequences of the type II *oriT*, *traM*, and *traY* alleles from ColB4-K98 and the type IV *traY* allele from R100-1. *J. Bacteriol.* **168**: 132-139.

Finlay, B.B., Frost, L.S., and Paranchych, W. (1986b) Nucleotide Sequence of the *traYALE* Region from IncFV Plasmid pED208. *J. Bacteriol.* **168**: 990-998.

Finlay, B.B., and Paranchych, W. (1986) Nucleotide sequence of the surface exclusion genes *traS* and *traT* from the IncF0 lac plasmid pED208. *J. Bacteriol.* **166**: 713-721.

Finlay, B.B., Paranchych, W., and Falkow, S. (1983) Characterization of conjugative plasmid EDP208. *J. Bacteriol.* **156**: 230-235.

Finnegan, D., and Willetts, N.S. (1973) The Site of Action of the F Transfer Inhibitor. *Mol. Gen. Genet.* **127**: 307-316.

Folkhard, W., Leonard, K.R., Malsey, S., Marvin, D.A., Dubochet, J., Engel, A., Achtman, M., Helmuth, R. (1979) X-ray diffraction and electron microscope studies on the structure of bacterial F pili. *J. Mol. Biol.* **130**: 145-160.

Freifelder, D. (1968) Studies with *E. coli* sex factors. Cold Spring Harbor Sym. *Quant. Biol* **33**: 425-434.

Friedman, D.I. (1988) Integration Host Factor: A Protein for All Reasons. *Cell* **55**: 545-554.

Frost, L.S., Finlay, B.B., Opgenorth, A., Paranchych, W., Lee, J.S. (1985) Characterization and Sequence Analysis of Pilin from F-Like Plasmids. *J. Bacteriol.* **164**: 1238-1247.

Frost, L.S., Paranchych, W., and Willetts, N.S. (1984) DNA sequence of the F *traALE* region that includes the gene for F pilin. *J. Bacteriol.* **160**: 395-401.

Fu, F.Y.-H., Tsai, M.-M., Luo, Y., Deonier, R.C. (1991) Deletion analysis of the F plasmid *oriT* locus. *J. Bacteriol.* **173**: 1012-1020.

Gaffney, D., Skurray, R., and Willetts, N. (1983) Regulation of the F conjugation genes studied by hybridization and *tra-lacZ* fusion. *J. Mol. Biol.* **168**: 103-122.

Gamas, P., Caro, L., Galas, D., Chandler, M. (1987) Expression of F transfer functions depends on the *E.coli* integration host factor. *Mol. Gen. Genet.* **207**: 302-305.

Guiney, D.G. (1982) Host range of conjugation and replication functions of the *Escherichia coli* sex plasmid F'*lac*: Comparison with the broad-host-range plasmid RK2. *J. Mol. Biol.* **162**: 699-703.

Harrington, L.C., and Rogerson, A.C. (1990) The F pilus of *Escherichia coli* appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. *J. Bacteriol.* **172**: 7263-7264.

Hayes, W. (1953) Observations on a transmissable agent determining sexual differentiation in *E. coli*. *J. Gen. Microbiol.* **8**: 72-88.

Heinemann, J.A., and Sprague, G.F.,Jr. (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340**: 205-209.

Inamoto, S., Abo, T., and Ohtsubo, E. (1990) Binding sites of Integration Host Factor in *oriT* of plasmid R100. *J. Gen. Appl. Microbiol.* **36**: 287-293.

Inamoto, S., Yoshioka, Y., and Ohtsubo, E. (1991) Site- and strand-specific nicking *in vitro* at *oriT* by the TraY-TraI endonuclease of plasmid R100. *J Biol Chem.* **266**: 10086-10092.

Ippen-Ihler, K., Moore, D., Laine, S., Johnson, D.A., Willetts, N.S. (1984) Synthesis of F-pilin polypeptide in the absence of F *traJ* product. *Plasmid.* **11**: 116-129.

Ippen-Ihler, K., and Maneewannakul, S. (1991) Conjugation among enteric bacteria: mating systems dependent on expression of pili. In: *Microbial Cell-Cell Interactions*, edited by Dworkin, M.,Washington, D.C.: American Society for Microbiology, 35-69.

Ippen-Ihler, K.A., and Minkley, E.G.,Jr. (1986) The conjugation system of F, the fertility factor of *Eschericia coli. Annu. Rev. Genet.* **20**: 593-624.

Jacob, F., and Wollman, E.L. (1961) *Sexuality and the genetics of bacteria*, New York:Academic Press, 374-384.

Kennedy, N., Beutin, L., Achtman, M., Skurray, R. (1977) Conjugation proteins encoded by the F sex factor. *Nature* **270**: 580-585.

Kingsman, A., and Willetts, N. (1978) The requirements for conjugal DNA synthesis in the donor strain during F*lac* transfer. *J. Mol. Biol.* **122**: 287-300.

Kolodkin, A.L., Capage, M.A., Golub, E.I., Low, K.B. (1983) F sex factor of *Escherichia coli* K12 codes for a single-stranded DNA binding protein. *Proc. Natl. Acad. Sci. USA* **80**: 4422-4426.

Kuhn, B., Abdel-Monem, M., Krell, H., Hoffman-Berling, H. (1979) Evidence for two mechanisms for unwinding catalyzed by DNA helicases. *J. Biol. Chem.* **254**: 11343-11350.

Lahue, E.E., and Matson, S.W. (1988) *Escherichia coli* DNA helicase I catalyzes a unidirectional and highly processive unwinding reaction. *J. Biol. Chem.* **263**: 3208-3215.

Lahue, E.E., and Matson, S.W. (1990) Purified *E.coli* F-factor TraY protein binds *oriT*. *J. Bacteriol.* **172**: 1385-1391.

Lederberg, J., and Tatum, E. (1946) Gene recombination in *E. coli*. *Nature* **158**: 558.

Manning, P.A., and Achtman, M. (1979) Cell to cell interactions in conjugating *Escherichia coli*: the involvement of the cell envelope. In: *Bacterial outer membrane biogenesis and functions*, edited by Inouye, M. New York: Wiley, 409-447.

Manning, P.A., Morelli, G., Achtman, M. (1981) TraG protein of the F sex factor of *Escherichia coli* K12 and its role in conjugation. *Proc. Natl. Acad. Sci. U. S. A.* **78**: 7487-7491.

Marmur, J., Rownd, R., Falkow, S., Baron, L.S., Schildkrant, C., Doty, P. (1961) The nature of intergeneric episomal infection. *Proc. Natl. Acad. Sci. U. S. A.* **47**: 972-979.

Matson, S.W., and Morton, B.S. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J Biol Chem.* **266**: 16232-16237.

Mazodier, P., and Davies, J. (1991) Gene transfer between distantly related bacteria. *Ann. Rev. Genet.* **25**: 147-171.

Minkley, E.G.,Jr., Polen, S., Brinton, C.C.,Jr., Ippen-Ihler, K. (1976) Identification of the structural gene for F-pilin. *J. Mol. Biol.* **108**: 111-121.

Mullineaux, P., and Willetts, N. (1985) Promoters in the transfer region of plasmid F. *Basic. Life. Sci.* **30**: 605-614.

Novick, R.P. (1987) Plasmid Incompatibility. *Microbiol. Rev.* **51**: 381-395.

Panicker, M.M., and Minkley, E.G.Jr. (1985) DNA transfer occurs during a cell surface contact stage of F sex factor-mediated bacterial conjugation. *J. Bacteriol.* **162**: 584-590.

Perumal, N., and Minkley, E.G.,Jr. (1984) The product of the F sex factor *traT* surface exclusion gene is a lipoprotein. *J. Biol. Chem.* **259**: 5357-5360.

Rees, C.E.D., and Wilkins, B.M. (1989) Transfer of *tra* proteins into the recipient cell during bacterial conjugation mediated by plasmid CoIIb-P9. *J. Bacteriol.* **171**: 3152-3157.

Rees, C.E.D., and Wilkins, B.M. (1990) Protein transfer into the recipient cell during bacterial conjugation: studies with F and RP4. *Mol. Microbiol.* **4**: 1199-1205.

Reygers, U., Wessel, R., Müller, H., Hoffmann-Berling, H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J* **10**: 2689-2694.

Riede, I., and Eschbach, M.L. (1986) Evidence that TraT interacts with OmpA of *Escherichia coli*. *FEBS Lett.* **205**: 241-245.

Ross, W., Landy, A., Kikuchi, Y., Nash, H. (1979) Interaction of Int protein with Specific sites on λ *att* DNA. *Cell* **18:** 297-307.

Schwab, M., Gruber, H., and Hogenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* 5: 439-446.

Silverman, P.M., Wickersham, E., and Harris, R. (1991) Regulation of the F plasmid *traY* promoter in *Escherichia coli* by host and plasmid factors. *J. Mol. Biol.* **218:** 119-128.

Taylor, D.E. (1989) General properties of resistance plasmids. In: *Handbook of Experimental Pharmacology, vol. 91*, edited by Bryan, L.E. Berlin : Springer-Verlag, 325-357.

Thompson, R., and Taylor, L. (1982) Promoter Mapping and DNA Sequencing of the F Plasmid Transfer Genes *traM* and *traJ*. *Mol. Gen. Genet.* **188**: 513-518.

Thompson, T.L., Centola, M.B., and Deonier, R.C. (1989) Location of the nick at *oriT* of the F plasmid. *J. Mol. Biol.* **207**: 505-512.

Traxler, B.A., and Minkley, E.D. (1988) Evidence that DNA helicase 1 and *oriT* site-specific nicking are both functions of the F TraI protein. *J. Mol. Biol.* **204**: 205-209.

Traxler, B.A., and Minkley, E.G.Jr. (1987) Revised genetic map of the distal end of the F transfer operon: Implications for DNA helicase 1, nicking at *oriT*, and conjugal DNA transport. *J. Bacteriol.* **169**: 3251-3259.

Tsai, M-M., Fu, Y.H.F., and Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *Bacteriol.*.**172**: 4603-4609.

Wilkins, B.M., and Hollom, S.E. (1974) Conjugative synthesis of F*lac*+ and ColI DNA in the presence of rifampicin and in *Escherichia coli* K12 mutants defective in DNA synthesis. *Mol. Gen. Genet.* **134**: 143-156.

Willetts, N. (1972) Location of the origin of transfer of the sex factor F. *J. Bacteriol.* **112**: 773-778.

Willetts, N. (1977) The transcriptional control of fertility in F-like plasmids. *J. Mol. Biol.* **112**: 141-148.

Willetts, N., and Maule, J. (1986) Specificities of IncF plasmid conjugation genes. *Genet. Res.* **47**: 1-11.

Willetts, N., and Skurray, R. (1980) The conjugation system of F-like plasmids. Annu. *Rev. Genet.* **14**: 41-76.

Willetts, N., and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48**: 24-41.

Willetts, N.S. and Skurray, R. (1987) Structure and function of the F factor and mechanism of conjugation. In: *Escherichia coli and Salmonella typhimurium: Cellular and molecular biology*, edited by Neidhardt, F.C., Ingraham, J.B., Low, K.B., Magasanik, B., Schaecter, M. and Umbarger, H.E. Washington, D.C.: American Society for Microbiology, 1110-1133.

Yang, C.-C., and Nash, H.A. (1989) The interaction of *E.coli* IHF protein with its specific binding sites. *Cell* **57**: 869-880.

Zambryski, P. (1989) *Agrobacterium*-plant cell DNA transfer. In: *Mobile DNA*, edited by Berg, D.E. and Howe, M. M. Washington, D.C.: American Society for Microbiology, 309-333.

CHAPTER II

Characterization of the *oriT* Region of the IncFV Plasmid pED208[1]

A. Introduction

The conjugative plasmid pED208 is derepressed for pilus formation and DNA transfer due to the insertion of an IS2 element near the transfer operon promoter in the parental plasmid F_0*lac* (Finlay *et al.*, 1986c). The origin of transfer (*oriT*) is immediately upstream from the IS2 element on a 23 kb *Hin*dIII fragment (Finlay *et al.*, 1983). Sequence analysis of the region downstream from this insertion element revealed coding regions for proteins that were clearly related to the *traYALE* genes of the IncF plasmids and the *traT* surface exclusion protein of the F and R100 plasmids (Finlay and Paranchych, 1986). Thus the variation in sequence at the DNA level belies the high degree of homology in the protein sequence and gene order between pED208 and other F-like plasmids.

The origin of transfer in conjugative plasmids is the site where one strand of the DNA is nicked. This is followed by the transfer of the nicked strand in a 5' to 3' direction into the recipient cell through a putative conjugational bridge. The nicking reaction involves the *traY* gene product (Willetts and Skurray, 1986; Ippen-Ihler and Minkley, 1986), and TraI which is DNA helicase I (Traxler and Minkley, 1988).

[1] A version of this chapter has been published. Di Laurenzio, L., Frost, L. S., Finlay, B. B., and Paranchych W. Mol. Microbiology. 1991. **5**: 1779-1790.

The TraM protein is thought to process the signal that a competent mating pair has formed and to initiate the transfer process (Thompson and Taylor, 1982; Willetts and Skurray, 1986). In addition to the nicking site, the *oriT* region contains a number of direct and inverted repeats which may be control elements in the transfer process. DNA sequence analysis of the *oriT* regions of the prototype plasmids that define the various classes of *oriT* revealed that the *oriT* region could be divided into three sections; the nicking region, an AT-rich region thought to bind the TraYI proteins and a region with a complex pattern of direct and inverted repeats which could be the binding site for the TraM protein. The number of different sequences in each of these regions corresponds to the number of homologs for *traI*, *traY* and *traM* (Finlay *et al.*, 1986b).

In addition to the *oriT* region and *traM*, the sequence presented in this Chapter revealed an open reading frame which is homologous to the TraJ protein of other IncF plasmids (Fowler *et al.*, 1983; Paranchych *et al.*, 1986). The TraJ protein is thought to be a positive regulatory element in the transcription of the transfer operon which encodes the genes for pilus assembly. However, no direct evidence of its binding to the promoter region has been observed (reviewed in Ippen-Ihler and Minkley, 1986; Paranchych and Frost, 1988). In the course of investigating *traM*, the *traJ* gene was sequenced and the composition and amino-terminal of the TraJ protein sequence were confirmed. However,understanding the mode of action of the TraJ protein requires further study.

B. Experimental Procedures

1. Bacterial strains and plasmids

pED208 is a derepressed mutant of the IncFV plasmid F_o*lac* (Finlay *et al.*, 1984). Recombinant plasmids, described below, were transformed into *E. coli* JM83 (Vieira and Messing, 1982), HB101 (Boyer and Roulland-Dussoix, 1969), K38 (Russel and Model, 1984) or JC3272 (Achtman *et al.*, 1971). *Escherichia coli* ED24 (F^- Lac Spc) was used as the recipient strain in plasmid transfer experiments (Finnegan and Willetts, 1971). The plasmids pT7-3, pT7-4 and pGP1-2 (Tabor and Richardson, 1985) were provided by Stan Tabor (Department of Biological Chemistry, Harvard Medical School, Boston, Ma.). The plasmid pUC18 (Yanisch-Perron *et al.*, 1985) was employed for cloning experiments. Chimeric plasmids used in this study are listed in Table II. 1.

2. Media

LB (Luria-Bertani) broth was as described previously (Maniatis *et al.*, 1982). ^{35}S-labeling minimal medium (Miller, 1972) was supplemented with 0.004% amino acids and 1% Methionine Assay powder (Difco Laboratories, Detroit, Mich.). Antibiotic concentrations were 25 μg ml^{-1} for both ampicillin and kanamycin (stock solutions 5 mg ml^{-1} and 10 mg ml^{-1} in water, respectively) and 100 μg ml^{-1} for rifampin (stock solution 20 mg ml^{-1} in methanol) and were obtained from Sigma Chemical Co., St. Louis, Mo.

3. Transfer efficiency assay

Mating experiments were performed by Dr Laura S. Frost. Cultures of donor (*E. coli* JC3272 containing various plasmid constructs) and recipient (*E. coli* ED24) cells were grown to mid-log phase and 0.1 ml of each was mixed together with 1.0 ml fresh LB broth. The mating mixture was incubated without shaking for 30 min at 37°C and appropriate dilutions were plated out on minimal plates containing spectinomycin (10 µg ml^{-1}) and either 2% glucose or 1% lactose and selective antibiotics.

4. Protein over-expression and labeling

The gene products expressed by the pT7-3 or pT7-4 chimeras (described in Table II. 1) were labeled with [^{35}S]-methionine as follows. *E. coli* K38 cells, transformed with both pGP1-2 and pLDL002 or pLDL005 were grown in 2 ml of LB medium with kanamycin and ampicillin at 30°C, to an OD_{600} of 0.5. one ml of culture was centrifuged and washed with ^{35}S-labeling medium. The cells were resuspended in 1.0 ml of ^{35}S-labeling medium and incubated with shaking at 30°C for 1 hr. The temperature was raised to 42°C for 10 min. A 5 µl volume of rifampin was added, and the cells were incubated for a further 10 min at 42°C. The cells were shifted to 30°C for another 20 min which was followed by the addition of 10 µCi of [^{35}S]-methionine (Dupont, Lachine, Quebec, Canada). After 5 min at 30°C, the cells were centrifuged, lysed with sample buffer (see below) and subjected to polyacrylamide gel electrophoresis (PAGE) on a 15% gel in the presence of sodium dodecyl sulfate (SDS). The gels were soaked for 30 min in 1M sodium salicylate and vacuum-dried before exposure to XAR-5 X-ray film (Eastman Kodak Co., Rochester, New York). To over-produce TraM protein, *E. coli* K38 cells, transformed with pGP1-2 and pLDL002, were grown to an OD_{600} of 0.9 in 200 ml of LB broth in the presence of kanamycin and ampicillin at 30°C. The cells were incubated in a 65°C water bath with swirling until the temperature reached 42°C (about 3 min) and were further incubated for 25 min at 42°C before addition of 1.0 ml of rifampin. The cells were ready for fractionation after 2 hr at 37°C.

5. Cell fractionation and membrane preparation

After over-production of the proteins, the cells were cooled on ice and centrifuged at 10,000g for 10 min. The pellet was washed and resuspended in 2.4 ml of 50 mM Tris-HCl, pH 8, and 25% sucrose. A 16 µl volume of 12 mg/ml of lysozyme in 0.25 M EDTA was added (final concentration, 80 µg ml^{-1} and 1.7 mM, respectively). Two volumes of ice cold water were added and the cells were incubated on ice for 30 min. The spheroplasts were centrifuged for 10 min at 12,000g. The supernatant containing the periplasmic fraction was dialysed and freeze-dried. The pellet was washed with Tris HCl, pH 8, 25% sucrose, and resuspended in 3.5 ml of ice cold water. The cell suspension was then passed three times through a French pressure cell (American Instrument Co, Silver Spring, Md.) at 986 Kg cm^{-2}. The crude lysate was centrifuged at 12,000g for 15 min to remove undisrupted cells and the supernatant was centrifuged at 100,000g in an ultracentrifuge (Beckman model L8-M, SW41 rotor, Beckman Corp., Sunnyvale, Ca.). The supernatant, containing the cytoplasmic fraction, was dialysed and freeze-dried. The crude membrane pellet was resuspended by vortexing and sonication (3 pulses of 30 sec each) with a microprobe in 10 ml of 10% (w/v) sucrose, 3 mM EDTA, pH 7.5, and centrifuged at 100,000g for 1 hr. The procedure for washing the membrane fraction was repeated (Yamato *et al.*, 1975). The crude membrane fraction was resuspended in 3 ml of 0.5% N-lauroylsarcosine (Sarkosyl, Sigma Chemical Co., St Louis, Mo.), incubated for 20 min at room temperature and sucrose was added to a final concentration of 0.25 M (Filip *et al.*, 1973). The inner and outer membranes were separated by ultracentrifugation at 360,000g for 2 hr in a SW50.1 rotor. The supernatant, containing the inner membrane proteins, was dialysed and freeze-dried, and the pellet containing the outer membrane proteins was resuspended in 3 mM EDTA, pH 8. All operations were performed at 4 °C. The total protein content in the four fractions was assayed using the Lowry (Lowry *et al.*, 1951) or the Biorad (BioRad, Richmond, Ca.) protein assays.

6. Crude cell extract preparation

The procedure for preparation of crude cell extracts was identical to that for cell fractionation except for the following. After treatment with lysozyme, the cells were disrupted by two passages through a French pressure cell and the lysate was centrifuged at 100,000g for 3 hr at 4°C. The soluble fraction was extensively dialysed against the buffer used in DNA electrophoretic mobility retardation assays (see below) and frozen at -20°C in the presence of 0.1 mM phenylmethylsulphonyl fluoride (PMSF). One ml aliquots were stored at 4°C for several months and used in DNA electrophoretic mobility retardation assays and DNA footprinting experiments.

7. TraM protein purification

The procedure for the preparation of the soluble fraction (fraction I) was identical to that of crude cell extract preparation except for the following. After lysozyme treatment, the cells were disrupted by two passages through a French press, RNase was added to a final concentration of 50 $\mu g\ ml^{-1}$, and the lysate was centrifuged at 100,000g for 1 hr. All steps in the purification procedure were performed at 4°C unless otherwise stated. Solid ammonium sulfate was added to fraction I to a concentration of 30% with constant stirring and centrifuged at 10,000g for 30 min. The ammonium sulfate was increased to 50% for 30 min and the solution was centrifuged at 10,000g for 30 min. The final supernatant (Fraction III) was dialysed against 50 mM Tris-HCl pH 7.5, 0.1 mM EDTA, 1mM dithiothreitol (DTT, ICN Biochemical, Cleveland, Ohio) (TED buffer) and concentrated 10 times using an ultrafiltration cell with a molecular weight cut-off of 10 kDa (Filtron Technology Corp., Clinton, Ma.).

One milliliter aliquots were loaded on an FPLC anion exchange column (Preparative MonoQ HR10/10, Pharmacia, Uppsala, Sweden) equilibrated with buffer TED. The samples were eluted using a linear NaCl gradient in TED at the flow rate of 1 ml min^{-1}. One milliliter fractions were collected and the elution profile was followed with a 280nm UV detector. TraM was eluted as a single peak at 0.30 - 0.32 M NaCl as determined by conductivity measurements. Peak fractions from 3 chromatography experiments were pooled, dialysed against TED, and concentrated to 300 µl using a microconcentrator with a molecular weight cut-off of 30 kDa (Amicon, Danvers, Ma.) (Fraction IV). A 10 µl volume of Fraction IV was loaded onto a native DNA-cellulose (Pharmacia, Uppsala, Sweden) column (1.0 cm, 1.6 cm) equilibrated with TED buffer at 20°C. The column was washed with 1 ml volumes of TED of increasing NaCl concentration. Pure TraM was detected in the 0.3 M NaCl/TED fraction (Fraction V).

8. Antisera preparation

Purified TraM and TraJ proteins were electro-eluted from lightly stained bands on a sodium dodecyl sulfate (SDS) polyacrylamide gel (PAG) according to the method of Hunkapiller *et al.* (1983). To obtain anti-TraM or anti-TraJ antisera, New

Zealand White rabbits were injected intramuscularly in the gluteal area with a total of 80 µg of pure protein dissolved in equal volumes of sterile saline solution and Freund's complete adjuvant (2 ml total). The rabbits were injected one and two weeks later with 80 µg of purified protein and Freund's incomplete adjuvant. The animals were bled approximately two weeks after the final injection.

9. Transfer of protein from SDS-polyacrylamide gels to nitrocellulose

SDS-PAG electrophoresis was carried out as described by Lugtenberg *et al.* (1975) with the following modifications. The running gel contained 15% acrylamide, 0.27 N,N'-methylene-bis-acrylamide, whereas the stacking gel contained 7.5% acrylamide, 0.2% N,N'-methylene-bis-acrylamide. The thickness of the gels was 0.75 mm. Samples were boiled for 5 min in sample buffer (0.08 M Tris-HCl, pH 6.8, 2.5% SDS, 10% glycerol, 0.7 M β-mercaptoethanol, 0.025% bromophenol blue) before application to the gel. After electrophoresis, the proteins were transferred to nitrocellulose paper with a Transblot apparatus (Biorad, Richmond, Ca.) according to the procedure described by Towbin *et al.* (1979). The transfer process was carried out for 120 min. A duplicate gel was also stained with Coomassie blue, using the method of Fairbanks *et al.* (1971).

10. Immunological detection of proteins on nitrocellulose paper

Immunological detection was exactly as described previously by Towbin *et al.* (1979) except that ^{125}I-labeled protein A (Dupont, Lachine, Quebec, Canada) was used instead of ^{125}I-labeled sheep immunoglobulin G.

11. Protein composition and sequencing

Crude extracts of cells over-expressing pLDL002 were separated by SDS-PAGE (Laemmli, 1970). Gel slices containing TraM or TraJ protein were excised from the gel and the protein was electro-eluted according to Hunkapiller *et al.* (1983). Amino acid analyses were performed on 3-5 nmol of protein hydrolysed in constant boiling HCl at 110ºC for 20 hr under vacuum. After hydrolysis, the samples were freeze-dried and washed in water several times to eliminate excess ammonium ion. The analysis was performed by Mike Natriss on a Dionex amino acid analyser fitted with a standard ion exchange column (DC6A, sulfonated polystyrene). Stepwise elution was performed with the following buffers; 0.2 M Na citrate buffer, pH 3.28, 0.2 M Na citrate buffer, pH 4.25, 0.2 M Na citrate and 0.9 M NaCl , pH 7.9. The column was calibrated with Amino Acid Standard H (Pierce Chemical Co., Rockford, Ill.).

Amino acids were detected by post-column ninhydrin derivatization and the effluents were monitored at 590 nm for amino acid derivatives and at 440 nm for proline derivatives using 51ºC for the first 50 min and 65ºC for the remaining 70 min. The amino acid composition of purified TraM protein was determined on 0.3 nmol of protein hydrolysed at 110°C for 20, 44, and 72 hours. The analysis was performed on an Applied Biosystems Inc. HPLC system consisting of Model 130A Separation system, Model 420A Derivatizer for phenylthiocarbamyl (PTC-) derivative identification. Automated protein sequence analysis (Hewick *et al.*, 1981) was performed by Mike Carpenter on an Applied Biosystem Model 470A gas-phase sequencer coupled to an Applied Biosystems Model 120A HPLC for on-line PTC-derivative identification. All sequencer chemicals were purchased from Applied

Biosystems Inc., Foster City, Ca. and the manufacturer's standard sequencing protocol (03RPTH) was used.

12. Recombinant DNA techniques

Restriction enzymes, T4 DNA ligase, and the large fragment of DNA polymerase I (Klenow) were purchased from Bethesda Research Laboratories, Gaithersburg, Md.. The restriction enzymes and T4 ligase were used according to the supplier's specifications. Nick translations, DNA transfer to nitrocellulose filters, DNA filter hybridization and colony filter hybridization have been described previously (Frost *et al.*, 1985). The transformation of *E. coli* with plasmid DNA and the isolation of DNA fragments for cloning procedures were according to Maniatis *et al.* (1982).

13. Preparation of labeled fragments

The plasmids pLDL101 and pLDL102 were digested with *Hin*dIII and *Eco*RI and the plasmid pLDL108 was digested with *Bam*HI and *Kpn*I. The products were separated by electrophoresis on a 1.6% agarose gel. The appropriate fragments, containing the *oriT* region, were eluted from the gel according to the procedure of Dretzen *et al.* (1981). The purified fragments were specifically labeled at the 3' end with $\alpha[^{32}P]$dCTP or $\alpha[^{32}P]$dATP (Dupont, Lachine, Quebec, Canada) by the "filling-in" reaction (Maniatis *et al.*, 1982).

14. DNA-protein complex electrophoretic mobility retardation assay

Protein-DNA complexes were detected using the method described by Garner and Revzin (1981) with the following modifications. ^{32}P end-labeled fragments were incubated for 30 min at room temperature in a 20 µl volume of DNase I footprinting buffer (25 mM HEPES, pH 7.8, 50 mM KCl, 0.5 mM DTT, 0.05 mM EDTA, 5% glycerol, 0.5 mM PMSF) with varying amounts of purified TraM protein in the presence of competitor DNA (poly[dI-dC]poly[dI-dC], Sigma Chemical Co., St. Louis, Mo.). Reaction mixtures were immediately loaded on pre-run 5% polyacrylamide gels without addition of any loading buffer (the glycerol present in the reaction buffer was sufficient to ensure proper loading into gel wells) and electrophoresis was performed at 30 mA at room temperature. The dried gels were exposed to X-Ray film at -70°C.

15. DNase I footprinting

DNase I (Deoxyribonuclease I, Worthington Biochemicals, Freehold, N.J.) protection analysis was performed on ^{32}P end-labeled DNA fragments as previously described (Ohlsson and Edlund, 1986). The products of DNase I digestion were separated either on 6% or 12% polyacrylamide-urea sequencing gels and G and G+A sequencing reactions were performed according to Maxam and Gilbert (1980) and used as size markers. Radioactivity was detected by exposure of the dried gel to X-ray film.

C. Results

1. Sequence analysis of the *oriT* to *traY* region of pED208

The sequence of the first 1782 bp of a 2.2 kb *Eco*RI-*Hin*dIII fragment derived from pED208, beginning at the *Eco*RI site and continuing to the beginning of the IS2 insertion element at position 1783 (the *Hin*dIII site occurs in the IS2 element), is presented in Fig. II.1. It was determined by Brett Finlay and Laura Frost. The sequence includes the important restriction sites, the organization of the genes in this region which have high homology to genes in the previously reported *oriT* regions of other F-like plasmids, and the site of insertion of the IS2 element in the *traY* gene. Examination of the first 360 base pairs of sequence beyond the *Eco*RI site reveals three regions that have homology to the *oriT* region of other F-like plasmids (Finlay *et al.*, 1986b). The first region (Fig. II. 1, positions 50-70) resembles the AC-rich region complementary to the nick site proposed for these plasmids. The second region, positions 67-137, is AT-rich and contains imperfect direct repeats starting at positions 94 and 116 and an inverted repeat at position 65-81. This is similar to the proposed binding site for the other F-like TraY proteins and, recently, the TraY protein of R100 was shown to bind to short repeats in this region (Inamoto and Ohtsubo, 1990). The four other F-like *oriT* regions have little homology with each other beyond the AT-rich region up to the translational start site for *traM*. This third region could be the binding site for the TraM protein since the number of classes for the *traM* gene corresponds to the number of different sequences in this region. In pED208, there are two direct repeats starting at positions 152 and 197 and an inverted repeat starting at 232. The most remarkable aspect of this sequence (positions 153-359) is the presence of 14 *Hinf*I sites in this region, most having the sequence GAA/TTC as well as two additional *Hinf*I-like sites at positions 164 and 355. The palindromic nature of this sequence makes it possible to predict many overlapping direct and inverted repeats using computer algorithms. However, only these 2 direct repeats and the strong inverted repeat are indicated for simplicity.

Figure II. 1

Sequence of the first 1782 bp of pED208 after the *Eco*RI site. The major direct and inverted repeats are indicated by arrows. The *Hinf*I sites, including 2 *Hinf*I-like sites in the proposed Tra M-binding region between positions 164 and 355, are underlined and non-conserved nucleotides are notched (^). The principal restriction sites are indicated above the sequence. The sequence homologous to the nick site in the F plasmid is marked as <-nick->. The proposed CRP binding site is underlined. The start of the IS2 insertion element is shown as IS2->. RBS represents the predicted ribosome binding sites and M, J, and Y stand for the start of the *traM*, *traJ* and *traY* genes, respectively.

```
EcoRI    .         .        30         .         .        60         .         .        90         .         .   SspI120
GAATTCCTCCTGGCTGACATATTTCTGTATGCCTGCTCCTTAACATCAAAGTCCCACACCCGTCGCAAGATTTAATCTTGCATGATTATCCTTTATTTTCAGTGTATTACGTTAATATTT
                                                        <-nick->
         .       150         .         .       180      TaqI         .       210         .         .       240
TTATTGTTATTCTATGTCCGACATAGATACTAGATTCATATGCAAATCTTTAGTGAATCGTAAAGGAGTCGAATCCAGATTCATCTACGAATCAGACCACTAAATTAAGATAAATTAAAG

  Taql   .   Taql  .       270         .         .       300         .         .       330         .         .       360
ATTCGAATCTGGATTCGAATCCCTAATTTCGGTATAACAGATTCGCCTGTATGAAGTTGGTGAATCCAGACTGACTCTAATAAAATAAATGGCCGATTCACATGTGACTATTGATAATCC

    Dral         .       390 M       .         .       420         .      S3a.   S3a 450         .         .       480
GACAAATTTAAATTTCTAATAAGGTTTTTGAAATGCCAAAAATACAGACTTATGTGAACAATAATGTTTATGAACAGATCACCGATCTCGTTACCATCAGGAAACAGGAAGGTATAGAAG
                     RBS        MetProLysIleGlnThrTyrValAsnAsnAsnValTyrGluGlnIleLeuAspLeuValThrIleArgLysGlnGluGlyIleGluG
         .         .       510         .         .       540         .         .       570         .         .       600
AAGCATCGCTTTCAAATGTTTCTTCAATGCTGCTGGAACTAGGCCTGCGCGTCTATATGATTCAGCAGGAAAAGAGAGAGGGGGGCTTTAATCAGATGGAGTACAACAAGCTCATGCTGG
luAlaSerLeuSerAsnValSerSerMetLeuLeuGluLeuGlyLeuArgValTyrMetIleGlnGlnGluLysArgGluGlyGlyPheAsnGlnMetGluTyrAsnLysLeuMetLeuG
         .         .       630         .         .       660     Taql.         .       690         .         .       720
AAAACGTTTCACGGGTACGTGCTATGTGTACAGAAATTCTTAAAATGAGCGTCTTAAACCAGGAGTCGATTGCCAGCGGTAATTTTGATTACGCCGTTATCAAACCCGCTATTGATAAGT
luAsnValSerArgValArgAlaMetCysThrGluIleLeuLysMetSerValLeuAsnGlnGluSerIleAlaSerGlyAsnPheAspTyrAlaValIleLysProAlaIleAspLysP
         .         .       750         .     S3a .       780         .         .XmnI   810         .         .      DraI
TTGCCAGAGAGCAGGTCAGTATATTTTTCCCTGATGATGAGGATGATCAGGAGTGAACCTGGTTTCATTTCTAAAAGCCGGAATCATTTCCCTGATTCCGGCTTTTTTTTGGTCTCTTTA
heAlaArgGluGlnValSerIlePhePheProAspAspGluAspAspGlnGlu***
.        .         .       870         .         .       900         .         .       930         .         .       960
AAAATATGGAATCTTTACATTCGCGTTAGCACTTACGTCACATAAACCTGACACTTTGAATGTTCTGGTGGATTTGAAACGTCTGGTCGCTATTATTAATTTATCTAAGCCAAAAGAGGG

        DraI         .       990         .         .      1020         .         .      1050         .         . J    1080
TGTGAATAATTTAAAAGCCAGCAATGATAAATAACTCCACCATGTATTAAAAATTGCAATGGACCTGTTCATTGCTTGTGGGATTCTTTTGTCTTTTTTCAGGAGGCCTGGATGACTGTT
                                                                                                   RBS         MetThrVal
         .         .      1110         .         .      1140         .         .      1170         .         .      1200
GTTGACCCAGCACGTTTTATGTATGAACGTAATCATTTTCCATCTTTGACAGATAAAGAGTTTGAAACGTTAGTTTTATACTGTCAGATGATGAATGTTCAGATGGTTGCTGATTATCAG
ValAspProAlaArgPheMetTyrGluArgAsnHisPheProSerLeuThrAspLysGluPheGluThrLeuValLeuTyrCysGlnMetMetAsnValGlnMetValAlaAspTyrGln
         .         .      1230         .         .      1260         .         .      1290         .         .      1320
AATCGTAAGCCTGATGTAATCATTAAGCATCTGAAAAGTTGCAGACAAAAAATAGGTGTTGAGAGCGACTTTGAACTTTATTTCATAGTAATAAATAAGTTTGTTAACTTTGAGAGAGTT
AsnArgLysProAspValIleIleLysHisLeuLysSerCysArgGlnLysIleGlyValGluSerAspPheGluLeuTyrPheIleValIleAsnLysPheValAsnPheGluArgVal
         .         .      1350         .         .      1380         .         .      1410         .         .      1440
TTTCCAGAGTTAACATCAGAACAGATAAATATCCTGGCTGCATTTTCCTTCTATCCTAAAAGGTCAACAATTGCCCGGCGTTTTGATATATATCGGTGTGATATATATGATGAGCTAATT
PheProGluLeuThrSerGluGlnIleAsnIleLeuAlaAlaPheSerPheTyrProLysArgSerThrIleAlaArgArgPheAspIleTyrArgCysAspIleTyrAspGluLeuIle
         .         .      1470         .         .      1500  S3a    .         .      1530         .         .      1560
AAAATCAGGAACAACCTTGGTATTGAAGACTTGGAGTCTCTCAGGATGCTTTTTTTCATGAAGATCACAGTTTTTTTGTAAGTTTTAAGTGTTTATCTGAGTGTTTATTGGAGCTTGTTT
LysIleArgAsnAsnLeuGlyIleGluAspLeuGluSerLeuArgMetLeuPhePheMetLysIleThrValPheLeu***
         .         .      1590         .         .      1620         .         .      1650         .         .      1680
TTTTGATTTAGCTCAATCTTTATTCCATTGATTTTGTGAATTAAATCACCTTTTGAAAAATAATGTGTTGTTTTTTCGGTGAAAAATAAATTTACCCCAAACCCTTGAATGGAGGAAATT
                                        CRP
         .         .      1710         .      Y  .      1740         .         .      1770         .   IS2->
GACTCTATCTAATATCCCTATCGCAAATTGAAACAGGTAGCGAGAGATGAAGGAGCCAAAATCTAATATCAGAAAGCTAATTGACATTGGTGGGTTAATTGGTAGACT
                                     RBS      MetLysGluProLysSerAsnIleArgLysLeuIleAspIleGlyGlyLeuIleGly
```

As in the case of the other four classes of *traM*, there was no strong prediction for a promoter for the *traM* transcript (Finlay *et al.*, 1986b; Thompson and Taylor, 1982). The open reading frame that begins at position 393 reveals a protein sequence with high amino-acid sequence homology to the TraM protein sequences of other F-like plasmids. The sequence did not contain a strongly predicted domain for DNA binding with the helix-bend-helix motif (Takeda *et al.*, 1983) and it had a highly acidic carboxy-terminus with 6 acidic residues in the last 7 amino acids.

The region between *traM* and *traJ* in other F-like plasmids contains an inverted repeat which could be a rho-dependent termination site for the *traM* transcript and the transcriptional start site for the *traJ* gene (Paranchych *et al.*, 1986). The *traJ* transcript in the F-like plasmids has a 105 base untranslated region which is the site of action of the FinOP antisense RNA regulatory system (Finlay *et al.*, 1986a). pED208 has an inverted repeat that begins at position 793 which, on the basis of homology to the other F-like *traM* genes, could be a rho-dependent terminator for *traM* transcription. However, there are no other inverted repeats in this region which would be analogous to the stem-and-loop structures predicted for the FinOP regulatory system of the other F-like plasmids. There is an open reading frame that begins at position 1072, preceded by a strong ribosome binding site which encodes for a protein homologous to the TraJ proteins of the other F-like plasmids which are themselves only 21-23% homologous with each other (Paranchych *et al.*, 1986). The TraJ protein has a characteristically high content of aromatic amino acids (14%) and a predicted DNA binding domain in the carboxy-terminal portion of the protein which was also found in the other three alleles of *traJ* (Paranchych *et al.*, 1986). However, the pED208 TraJ protein is smaller than the other TraJ proteins, with a predicted molecular weight of 17,900. The sequence following the *traJ* gene could encode a longer polypeptide if there was a shift in the reading frame. The possibility that the *traJ* gene degenerated after the insertion of the IS2 element existed. Thus, the corresponding 3.075 kb *Eco*RI fragment of F_0*lac*, which encodes the *oriT* to *traE* genes was cloned and sequenced in the *traJ* region with a series of oligonucleotide primers. No difference between the *traJ* regions of pED208 and F_0*lac* could be detected; this suggested that the TraJ protein of pED208 is indeed smaller than those of the other F-like plasmids.

The region following the *traJ* gene contains two small direct repeats (positions 1527 and 1540), but no inverted repeats which have been proposed to be signals for termination of the *traJ* transcript or the initiation of the *traYI* operon in the other F-like plasmids (Paranchych *et al.*, 1986). However, the high AT content characteristic of this region is conserved in pED208 as well as a potential CRP (cAMP Receptor Protein) binding site (AA -TGTGA--T---TCA-TA/T, positions 1592-1612) (Ebright *et al.*, 1984).

An open reading frame begins at position 1727 which is preceded by a weak ribosome binding site and which encodes a peptide of 19 residues before the beginning of the sequence of the IS2 insertion element. When the sequences flanking the IS2 element presented here and in Finlay *et al.*, (1986c) are combined, there is a characteristic 5 bp repeated sequence (two codons repeated) on either side of the IS2 element (Kleckner, 1981) and there are 17 codons from the initiation site of *traY* to the beginning of the IS2 element. The sequence of the amino-terminal region has no discernible homology with that of the other three F-like TraY proteins

(Finlay *et al.*, 1986b). The F-like *traY* genes begin with the rare initiating codon, UUG, as well as a strong ribosome binding site (Inamoto *et al.*, 1988; Finlay *et al.*, 1986b). Examination of the region upstream of *traY* in pED208 revealed the presence of a strong ribosome binding site (positions 1671-1676) and a UUG codon three base pairs downstream, suggesting there is a primordial relationship to the other F-like plasmids in this region. However, there are 2 stop codons in this reading frame.

2. Determination of the minimal sequence required for transfer

Finlay *et al.* (1983) showed that cells containing a 30 kb *Hin*dIII fragment from pED208, beginning in the middle of the IS2 element and cloned into pACYC184 (pBF101), were capable of synthesizing pili but could not transfer. When a second plasmid (pBF105), containing a 23 kb *Hin*dIII fragment cloned into pBR322, was introduced into cells containing pBF101, the cells became transfer-proficient and pBF105 was transferred into the recipient cell. pLDL100 and its derivatives (Table II. 1) are subclones of pBF105. The first 70 base pairs downstream from the *Eco*RI site resemble the nicking site of other F-like plasmids. The ability of these plasmids to be mobilized in the presence of pBF101 was tested (see Experimental Procedures page 34). pLDL100 (containing the 2.2 kb *Eco*RI/*Hin*dIII fragment) and pLF201 (containing the 790 bp *Eco*RI/*Xmn*I fragment) were capable of being transferred while pLDL102 (containing the 190 bp *Eco*RI/*Taq*I fragment) was not (Table II. 1). Thus, *traM* and the proposed nicking region in the first 70 base pairs after the *Eco*RI site (Fig. II. 1) were required for transfer but sequences upstream of the *Eco*RI site as well as *traJ* and surrounding sequences were not essential.

Table II. 1

Plasmid	Portion of pED208*/Vector	Mobilization frequency per donor cell **
pLDL100	2.2 kb *Eco*RI-*Hind*III/pUC18	0.74
pLDL002	2.2 kb *Eco*RI-*Hind*III/pT7-3	ND
pLDL102	191 bp *Eco*RI-*Taq*I (1-189)/pUC18	ND
pLDL101	437 bp *Eco*RI-*Sau*3A(1-437)/pUC18	$<0.5 \cdot 10^{-4}$
pLDL108	253 bp *Ssp*I-*Dra*I (116-369)/pUC18	ND
pLDL005	470 bp *Dra*I-*Dra*I (369-839)/pT7-4	ND
pLF201	790 bp *Eco*RI-*Xmn*I (1-801)/pUC18	1.2
pBF101***	30 kb *Hind*III/pACYC184	0.7
pBF105***	23 kb *Hind*III/pBR322	0.0

* - Numbers in parentheses refer to position in Figure II. 1

**- The ability of a chimera to be mobilized into a recipient cell in the presence of pBF101, which is capable of pilus expression but lacks *oriT* and *TraMJ*, was measured as described in the Experimental Procedures

***- Finlay *et al.* (1983)

ND, not determined

3. Overproduction and analysis of the TraM and TraJ proteins

The 2.2 kb *Eco*RI-*Hin*dIII fragment was cloned into pT7-3 vector (Tabor and Richardson, 1985), named pLDL002 and expressed in *Escherichia coli* K38. A typical autoradiogram of [^{35}S]-methionine-labeled proteins from crude extracts of whole cells is shown in Figure II. 2. In addition to the β-lactamase protein and its precursor at 25-30 kDa, and an unknown protein band of 60-70 kDa, there were only two proteins expressed, which had apparent molecular weights of 10,900 and 15,300. The expected molecular weights for the TraM and TraJ proteins were 14,542 and 17,894, respectively, as calculated from the DNA sequence. No other coding regions in pLDL002 were capable of coding for a protein.

In order to obtain large amounts of the 10.9 and 15.3 kDa proteins, pLDL002 was over-expressed in *E. coli* K38 for a two hour period as described in Experimental Procedures. The cells were French-pressed and fractionated into soluble and membrane fractions, and the location of the putative TraM and TraJ proteins was determined by SDS-PAGE. The 10.9 kDa protein was found to be associated with the soluble cytoplasmic fraction while the 15.3 kDa band was found with the membrane fraction.

An SDS polyacrylamide gel of cell extracts was stained with Coomassie blue and the proteins of apparent molecular weights of 10,900 and 15,300 were electro-eluted according to the method of Hunkapiller *et al.* (1983). The amino acid composition and N-terminal sequence of the first 5 residues of each protein was determined and was in agreement with the predicted composition and sequence of

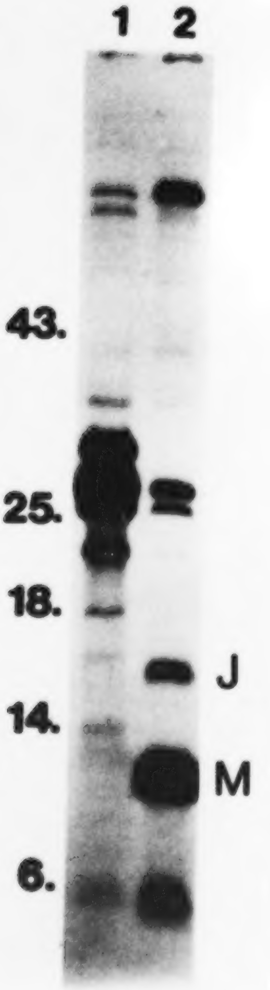


Figure II. 2

Autoradiography of an SDS gel of cells labeled with ^{35}S-methionine expressing TraM and TraJ proteins. 1; Control, pGP1-2 and pT7-3. 2; pGP1-2 and pLDL002. Molecular weight standards in kilodaltons are indicated on the left. M and J mark the positions of the 10.9 and 15.3 kilodaltons proteins, respectively.

TraM and TraJ. Thus, the 10.9 kDa and 15.3 kDa proteins were considered to be TraM and TraJ, respectively.

4. Immunological characterization of TraM and TraJ proteins

The electro-eluted TraM protein was also used to raise antibodies in rabbits as described in Experimental Procedures. The location of TraM in cells carrying the plasmid pED208 as well as the wild type plasmid F_0*lac* was determined by immunoblots of cells fractionated into cytoplasmic and periplasmic fractions and inner and outer membranes as described in Experimental Procedures. The proteins were separated on duplicate 15% SDS acrylamide gels and one was stained while the other was electrophoretically transferred to nitrocellulose. The immunoblot was treated with anti-TraM antiserum diluted 100-fold (Fig. II. 3). The TraM protein was found predominantly in the cytoplasm with smaller amounts in the periplasm and inner membrane. The periplasmic fraction was slightly contaminated with cytoplasmic proteins and the presence of TraM in the periplasm could be due to leakage from the cytoplasm during preparation. The presence of TraM protein in the inner membrane could represent a real association of TraM with one or more cellular or transfer proteins expressed by the plasmid. The anti-TraM antibodies reacted with unknown material of low molecular weight in the outer membrane which was also present in control samples prepared from plasmid free cells (data not shown).

5. TraM protein purification

In order to purify the TraM protein, the *Dra*I fragment between nucleotides 369 and 839, containing the *traM* gene, was subcloned into the expression vector pT7-4 (pLDL005). Two liters of *E. coli* JC3272 carrying pGP1-2 and pLDL005 were

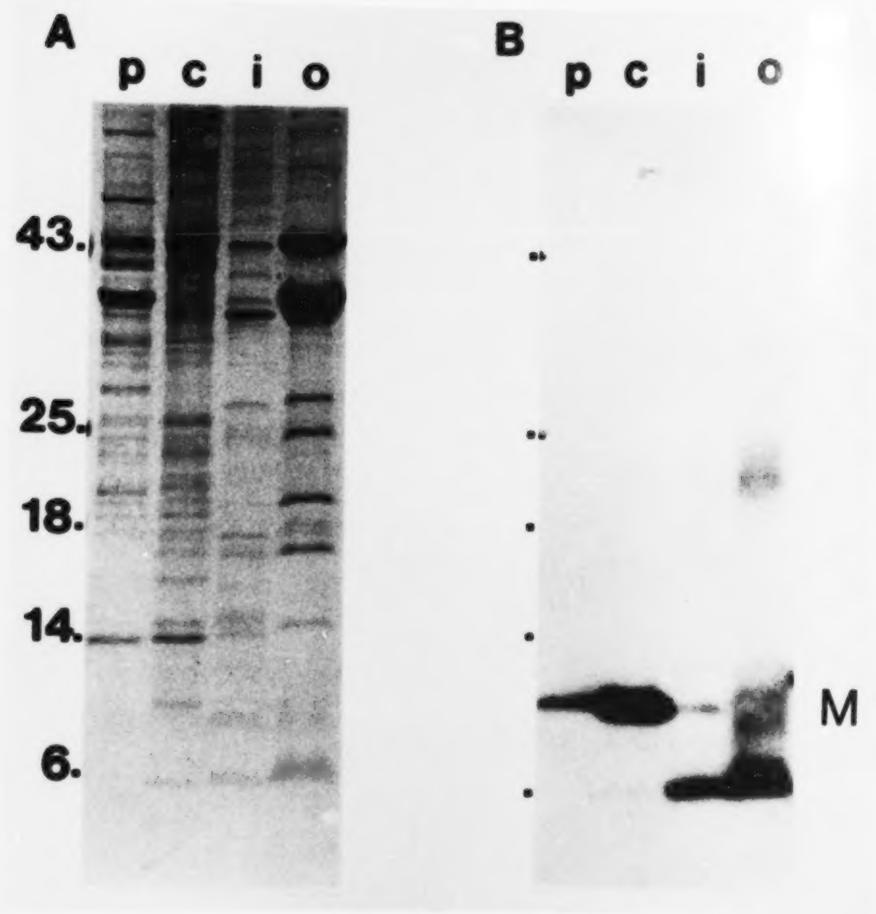


Figure II. 3

Localization of TraM protein in *E. coli* JC3272/pED208. Cellular fractions were prepared as described in Experimental Procedures. Approximately 30 μg of protein from each fraction were separated by SDS-PAGE. A. Coomassie blue stained gel. B. Immunoblot of a duplicate gel using anti-TraM antiserum. The molecular weight standards in kilodaltons are marked on the left and the position of TraM (M) is marked on the right. **p**, periplasm; **c**, cytoplasm; **i**, inner; and **o**, outer membrane fractions.

grown to an optical density of 1.1 at 600 nm and were incubated for a period of 150 minutes. The cells were passed through a French press twice and fractionated into soluble and membrane fractions (fraction I and II respectively). The TraM protein was purified from fraction I as described in Experimental Procedures and was monitored by SDS-PAGE followed by Coomassie blue staining (Fig. II. 4). When silver staining of a duplicate gel containing 5 μg of fraction V was performed, no extra bands appeared beside TraM (data not shown).

6. DNA electrophoretic mobility retardation assays

The ability of the TraM protein to bind to the *oriT* region of pED208 was determined using DNA electrophoretic mobility retardation assays. ^{32}P-labeled DNA fragments of various sizes were treated briefly with purified TraM and separated on 5% non-denaturing polyacrylamide gels. When the 370 bp *Eco*RI-*Dra*I fragment, containing the whole *oriT* region was used, a broad band of decreasing mobility was observed on the gel as the concentration of pure TraM increased (Fig. II. 5). The DNA-protein complexes resolved into a sharp band of high molecular weight at high TraM concentrations. We were not able to detect DNA in both the bound and unbound state in the same sample.

7. DNA footprinting of the *oriT* region with TraM protein

DNase I protection experiments were performed using a variety of ^{32}P-labeled DNA fragments derived from the *oriT* region of pED208 in order to assess the exact position of TraM protein binding. When the 189 bp *Eco*RI-*Taq*I fragment from pLDL101 was pretreated with TraM-containing crude extract, subjected to limited digestion by DNase I, and electrophoresed on a 12% sequencing gel, a single large

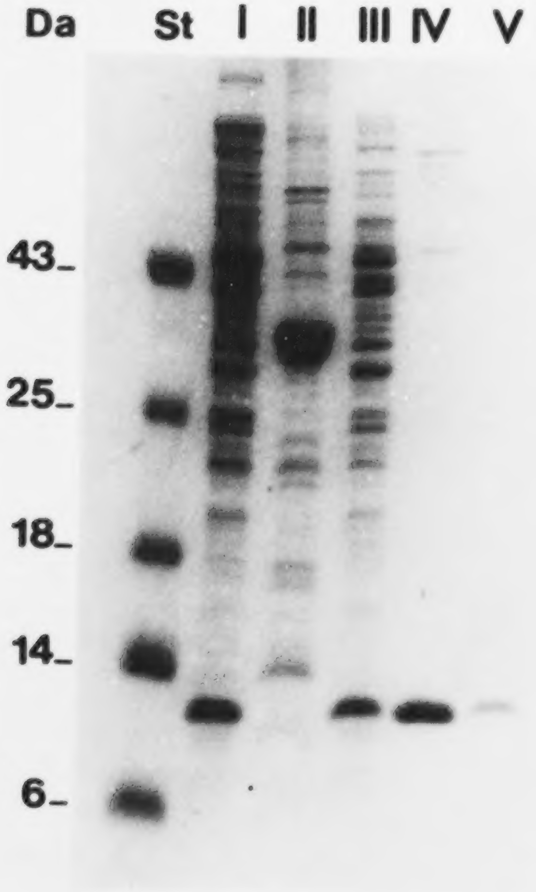


Figure II. 4

SDS-polyacrylamide gel (15%) electrophoresis of samples taken from the steps in the purification of TraM and stained with Coomassie blue. Lane I, 12.8 µg of soluble fraction (fraction I) ; lane II, 10 µg of crude membrane (fraction II); lane III, 11.2 µg of 50% ammonium sulfate (fraction III); lane IV, 5 µg of the peak eluted from the FPLC anion exchange column (fraction IV); lane V, 1 µg of the material eluted from the DNA-cellulose column (fraction V); St, molecular weight standards in kilodaltons.

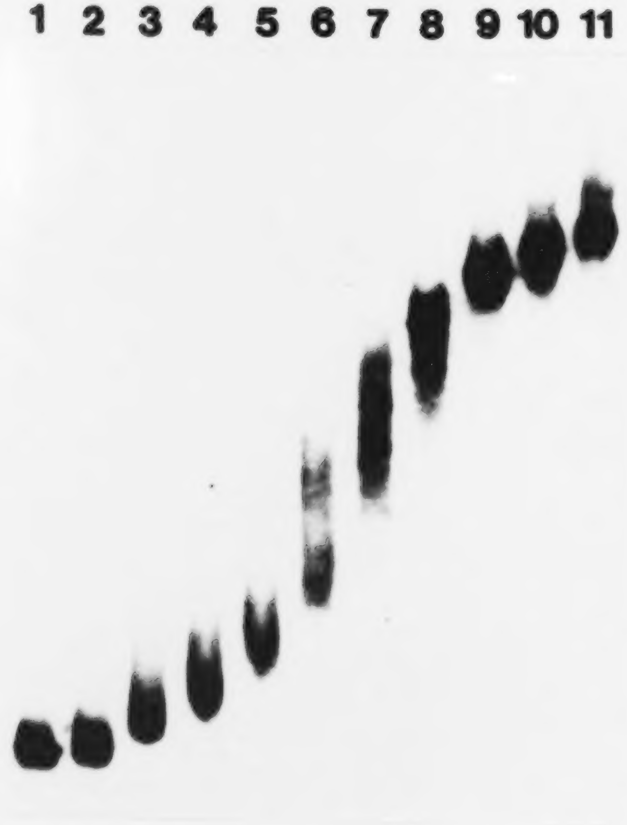


Figure II. 5

Binding of TraM protein to the *oriT* region. 2.70 fmol of the ^{32}P-labeled 370 bp *Eco*RI-*Dra*I fragment from pLDL101 was incubated with increasing amounts of purified TraM protein, in the presence of 1 μg of competitor DNA as described in Experimental Procedures. Lane 1 represents the free fragment. Lanes 2-11; Lane 1 with 1.5 ng, 3.0 ng, 9.0 ng, 21.0 ng, 30 ng, 90 ng, 0.3 μg, 1.5 μg, 3 μg, 9 μg, of purified TraM respectively.

footprint was obtained from positions 151 to the *Taq*I site at position 189 (Fig. II. 6, left panel). This suggested that there was no binding site for TraM in the first 150 base pairs. Overexposure of the gel revealed bands between the *Hinf*I-like sequences which occur at positions 153, 164, 175 and 186. An exception was the presence of a footprint covering the sequence AAAGGAG at positions 182-188. Thus, the large footprint could be broken down into many small footprints, each corresponding to a *Hinf*I-like sequence with an extra binding site for TraM or another protein at position 182-188. Crude extracts containing no TraM protein did not give a footprint at this position suggesting that the footprint was caused by TraM or some cellular protein acting in the presence of TraM.

In subsequent experiments, using either the *Eco*RI-*Hin*dIII fragment (data not shown) or the*Ssp*I-*Hin*dIII fragment (Fig. II. 6, right panel), both from pLDL101, crude extracts containing the TraM protein protected three regions in the *oriT* sequence. The most prominent footprint (I) was found to cover the inverted repeat that started at position 232, followed by a footprint (II) covering the two direct repeats at positions 152 and 197 and a third footprint (III) was observed in the region from 296 to 357 (Fig. II. 6). TraM protein appeared to bind to the footprints in the order I, II, III as increasing amounts of TraM were used (data not shown). The boundaries of the footprints were difficult to determine because DNase I has a preference for GC base pairs while this region is AT rich. However, if a number of gels were examined after lengthy exposures to X-ray film, smaller footprints corresponding to all the *Hinf*I-like sites except the one at position 280 could be seen.

Figure II. 6

DNase I protection of the *oriT* region in the presence of TraM. Left: the ^{32}P-labeled *Eco*RI-*Hin*dIII fragment from pLDL102 containing the 189 bp *Eco*RI-*Taq*I fragment, labeled at the 3'-end, in the presence of increasing amounts of crude extract containing TraM (Lanes 1-4; 0, 6.5, 13 and 32.5 µg of crude extract). TaqI indicates the position of the *Taq*I site at position 189. * indicates the position of *Hinf*I-like sites; (*) indicates a fortuitous *Hinf*I site (GAGTC) that arose from the cloning procedure. Right: the ^{32}P-labeled *Ssp*I-*Hin*dIII fragment from pLDL100, labeled at the 3' end, incubated with 18 and 27 µg of control crude extract over-expressing pT7-3 (lane 4 and 5), and 9, 18 and 27 µg of crude extract over-expressing the TraM protein (lanes 6 - 7 and 8) as described in Experimental Procedures. Lanes 1 and 2 are G+A and G sequencing reactions. Lane 3 is the DNA fragment incubated with DNase I in the absence of protein extracts. * indicates the position of the 16 *Hinf*I-like sites in the fragment. The Roman numerals indicate the position of the three large footprints as discussed in the text.

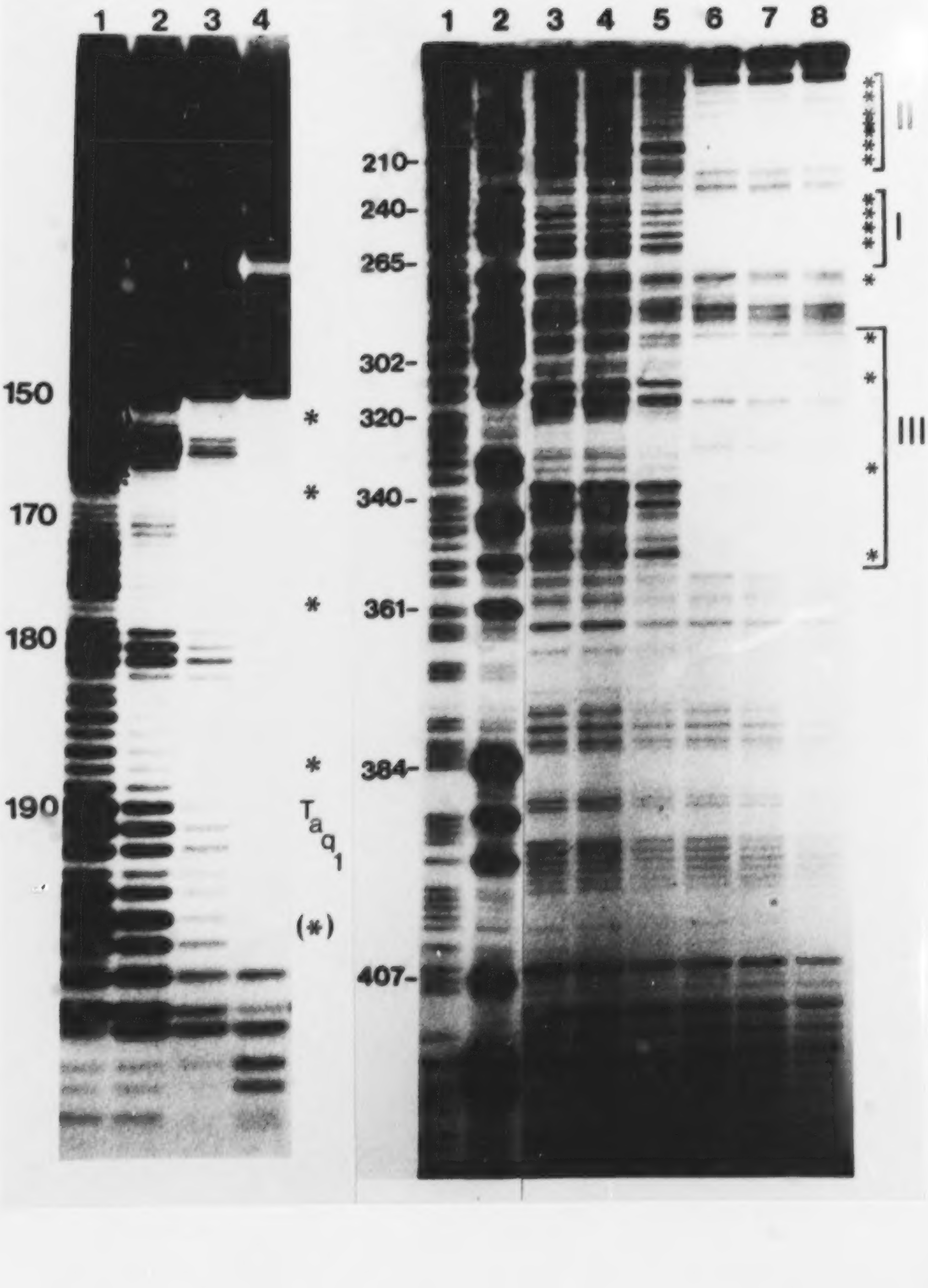

1 2 3 4
150
170
180
190
*
*
*
*
Taq1
(*)
1 2 3 4 5 6 7 8
210-
240-
265-
302-
320-
340-
361-
384-
407-
II
I
III
*

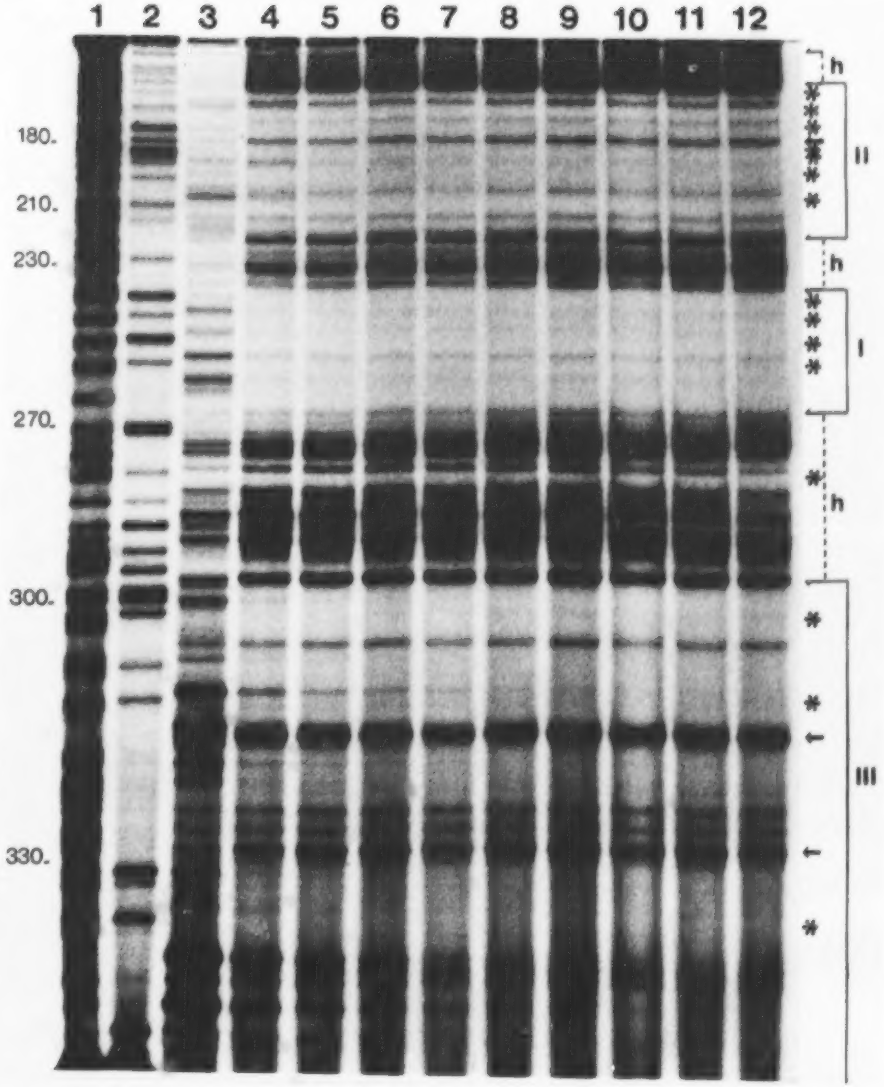
Figure II. 6 (right panel) also includes a control experiment performed using a cell extract containing pT7-3 instead of pLDL002. No footprint was obtained, suggesting that no cellular proteins bound the *oriT* region in the absence of the pED208 plasmid.

DNA footprinting experiments performed on the *oriT* region with purified TraM showed three protected regions delimited by hypersensitive sites. These regions were almost identical to the ones found using crude extract of cells over-expressing TraM protein. The main difference between the two sets of experiments was the presence of hypersensitive sites within region II and III at positions 180, 317, 329 (Fig. II. 7) and the lack of protection for the sequence AAAGGAG at position 182-188.

The *oriT* region is characterized by the presence of 14 *Hinf*I sites, 12 with the sequence GAA/TTC. Another two imperfect *Hinf*I-like sequences, with a one non-conserved base can also be found in the sequence from positions 153 to 355 (Fig. II. 1). All of these sites, except the one at 280, bound TraM protein. Analysis of the spacing of these sequences showed that most of them occur at 11-12 bp intervals or multiples thereof (Fig. II. 8). The inverted repeat (footprint I) contained two overlapping repeats 12 base pairs apart while footprint III (*traM* promoter region) contained these sequences at a distance of 11 and 21-22 base pairs apart. If the binding site has the sequence GATTC, then TraM would bind to both strands of the DNA in a scattered fashion. However, if the binding site has the sequence GANTC, then TraM protein would bind to as many as 30 sites on the DNA and the entire region would be coated with TraM protein.

Figure II. 7

DNase I protection of the *oriT* region in the presence of purified TraM. ^{32}P- labeled *Bam*HI-*Kpn*I fragment from pLDL108, containing the 253 bp *SSp*I-*Dra*I fragment, labeled at the 3' end, incubated with increasing amounts of purified TraM (lanes 4 to 12 contain between 0.15-15 µg of protein). Lane 3, DNA fragment digested with DNase I in the absence of protein; lanes 1 and 2, G+A and G sequencing reactions. * indicates the positions of the *Hinf*I-like sites. The arrows indicate the hypersensitive sites at position 180, 317 and 329. Hypersensitive regions (h) are marked by dashed lines whereas the protected regions are bracketed and numbered with Roman numerals as discussed in the text. The first 50 bp of the DNA were run off in order to have a better resolution of footprints I and II. A 12% sequencing gel of the same reactions was run to determine the boundaries of footprint III, which corresponded to the ones in Figure II. 6.


1
2
3
4
5
6
7
8
9
10
11
12
180.
210.
230.
270.
300.
330.
h
II
h
I
h
III
*
←

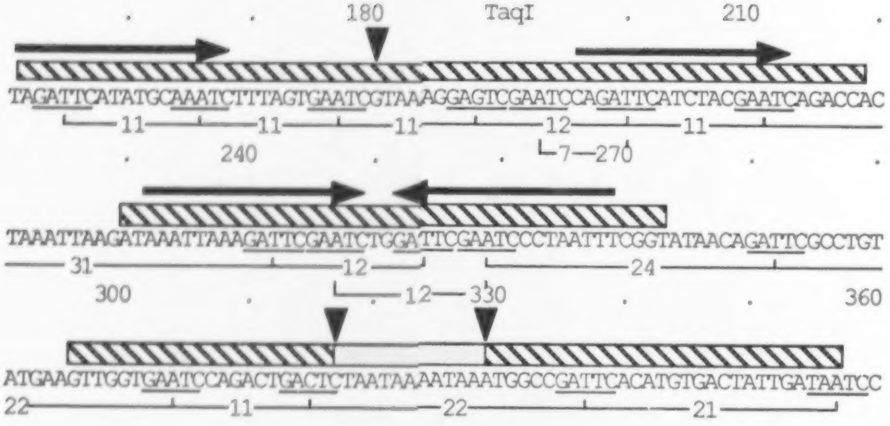

180
TaqI
210
TAGATTCATATGCAAATCTTTAGTGAATCGTAAAGGAGTCGAATCCAGATTCATCTACGAATCAGACCAC
11
11
11
12
11
240
7
270
TAAATTAAGATAAATTAAAGATTCGAATCTGGATTCGAATCCCTAATTTCGGTATAACAGATTCGCCTGT
31
12
24
300
12
330
360
ATGAAGTTGGTGAATCCAGACTGACTCTAATAAAATAAATGGCCGATTCACATGTGACTATTGATAATCC
22
11
22
21

Figure II. 8

Summary of the interactions between TraM protein and the 151-360 bp region of *oriT* of pED208. The *Hin*fI-like sites in the *oriT* region are underlined and the spacing between them is indicated by brackets. The direct and inverted repeats are marked with arrows. The regions protected by TraM during DNase I footprinting are indicated by striped boxes. The hypersensitive sites are marked by arrowheads and the partially protected region between them is indicated with a dotted box.

D. Discussion

The pED208 plasmid is a serendipitous mutant of an F-like plasmid, F_0*lac*, which confers multipiliation on its host and has its *oriT* region separated from the bulk of the transfer region by an IS2 element. The organization of its genes and the sequence of the genes studied to date show that it is homologous to the paradigmatic IncF plasmid, F. This includes the *oriT* region described in this Chapter. Because of the property of multipiliation, the odds of observing transfer-related processes are greatly enhanced because of a larger concentration of pilus-associated proteins. A derepressed mutant of F_0*lac* isolated in the laboratory of Dr. Paranchych (data not shown) produced 5-6 pili/cell which is within the normal range for F-like plasmids (Frost *et al.*, 1985).

The sequence presented here contains two prominent open reading frames which has homology to the TraM and TraJ proteins of other F-like plasmids. Their identity was confirmed by amino acid analysis and amino-terminal sequencing. When over-expressed from pLDL002, TraM was found in the soluble fraction while TraJ was found in the insoluble membrane fraction, allowing easy separation of the two proteins. Immunoblots of cells carrying pED208 or F_0*lac* indicated that TraM was found in the cytoplasm and inner membrane, which is consistent with its proposed role as a signalling protein in the initiation of transfer. Preliminary results suggested that TraJ was also found in the cytoplasm and membrane which may suggest a role for it as a sensor/regulator protein analogous to the "two component" regulatory systems (Ronson *et al.*, 1987) which sense changes in environmental conditions. However, the F-like TraJ proteins have no obvious sequence homology to this class of proteins. Cuozzo and Silverman (1986) demonstrated that the F TraJ protein, when expressed from the single copy F plasmid, was located in the cytoplasm and that over-expression from multi-copy plasmids or expression vectors caused it to aggregate into an insoluble mass which associated with the membrane. Either the pED208 TraJ protein, when expressed at normal levels, is extremely insoluble, or it acts in a different manner than the F TraJ protein. Efforts to demonstrate DNA binding of TraJ protein to fragments derived from the *traJ* to *traY* region (transfer operon promoter region) were unsuccessful and the mode of TraJ action remains elusive.

Electrophoretic mobility retardation assays using purified TraM protein incubated with a fragment containing the whole *oriT* region suggested that the binding of TraM to *oriT* was nonco-operative and that the TraM/*oriT* complex underwent exchange between binding sites during electrophoresis giving diffuse bands. Unlike other known DNA binding proteins, increasing concentrations of TraM do not bind to DNA fragments to give discreet steps representing the ordered loading of the binding sites. The anomalous mobility of DNA-protein complexes could be due to the bending of the DNA, rather than to the increased molecular weight of the complex compared to the free fragment. TraM binding at the *oriT* region occurs without marked changes in the DNA structure, as observed by electron microscopy (data not shown), and binding appears to be a gradual loading of many sites with similar affinity creating an heterogeneous mixture of complexes. The heterogeneity is lost once all the sites are bound.

The proposed site for TraM binding is different in sequence from those of other F-like plasmids (Finlay *et al.*, 1986b). However, R1 has five *Hinf*I sites evenly arranged in its proposed TraM protein binding region. pED208 TraM was found to bind weakly and in a competitor-sensitive manner to a fragment derived from the R1 *oriT* region but did not bind at all to a comparative fragment containing the F *oriT* region (data not shown). This suggests that pED208 and R1 may have evolved from a common ancestor with a similar *traM*/*oriT* locus.

DNA footprinting experiments showed that TraM protein binds to three principal regions in the *oriT* region, the inverted and direct repeats from positions 150 to 270 and the *traM* promoter region from positions 300 to 360 which included all the *Hinf*I-like sites except one at position 280. Regions II and III are divided into four subregions by the hypersensitive sites (see Fig. II. 8) containing three, four, two and two *Hinf*I-like sites, respectively. The difference between the footprints using pure TraM protein and crude extracts containing over-expressed TraM are striking in that the hypersensitive sites are not found using the crude extract. Similarly, there is an extra footprint at position 182-188 with crude extracts of TraM which is not present with the pure TraM protein nor with the control crude extract. This suggests that the TraM/*oriT* complex may induce binding of other host proteins, which cannot bind in the absence of TraM, and protect the DNA from hypersensitive attack by DNase I.

If the consensus sequence for TraM protein binding is GANTC, then TraM protein could bind up to 30 times in the *oriT* region. There is a half turn incorporated into the inverted repeat (two overlapping 12 base pair repeats, see Figure II. 8) which would allow TraM to bind both sides of the DNA either simultaneously or alternatively. The significance of this knot of TraM molecules within 22 bp of sequence (positions 240-261) is not readily apparent at this time.

We are currently investigating the nature of the association of the TraM protein with the inner membrane, be it alone or in a complex with other transfer and cellular proteins. Preliminary evidence suggests that the TraM protein alone sediments as a tetramer in sucrose gradients, while from cell extracts of pED208 it sediments as a large complex, suggesting its association with other proteins or DNA within the cell. The region that binds TraM protein is the last portion of the plasmid to enter the recipient cell (reviewed in Willetts and Skurray, 1986) and the potentially large number of TraM molecules bound to the DNA could resemble a knot on the end of a string and could be part of the signal that one round of plasmid transfer is complete. Also, the finding that TraM binds to its own promoter region, assuming that there is a *traM* promoter between *oriT* and *traM*, is intriguing and suggests that it may regulate its own expression.

E. Bibliography

Achtman, M., Willetts, N.S., and Clark, A.J. (1971) Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. *J. Bacteriol.* **106**: 529-538.

Boyer, H. W., and Roulland-Dussoix, D. (1969) A complementation analysis of the restriction and modification of DNA in *Escherichia coli. J. Med. Biol.* **41**: 459-472.

Cuozzo, M., and Silverman, P. M. (1986) Characterization of the F plasmid TraJ protein synthesized in F' and Hfr strains of *Escherichia coli* K-12. *J. Biol. Chem.* **261**: 5175-5179.

Dretzen, G., Bellard, M., Sassone-Corsi, P., and Chambon, P. (1981) A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. *Anal. Biochem.* **112**: 295-298

Ebright, R.H., Cossart, P., Gicquel-Sanzey, B., and Beckwith, J. (1984) Mutations that alter the DNA sequence specificity of the catabolite gene activator protein of *E. coli. Nature* **311**: 232-235.

Fairbanks, G., Steck, T. L., and Wallach, D. F. H. (1971) Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry* **10**: 2606-2617.

Filip, C., Fletcher, G., Wulff, J. L., and Earhart, C. F., (1973) Solubilization of the cytoplasmic membrane of *Escherichia coli* by the ionic detergent sodium-lauryl sarcosinate. *J. Bacteriol.* **115**: 717- 722.

Finlay, B.B., Paranchych, W., and Falkow, S. (1983) Characterization of conjugative plasmid EDP208. *J. Bacteriol.* **156**: 230-235.

Finlay, B. B., and Paranchych, W. (1986) Nucleotide sequence of the surface exclusion genes *traS* and *traT* from the IncF$_0$ *lac* plasmid pED208. *J. Bacteriol.* **166**: 713-721.

Finlay, B.B., Frost, L.S., Paranchych, W., and Willetts, N.S. (1986a) Nucleotide sequences of five IncF plasmid *finP* alleles. *J. Bacteriol.* **167**: 754-757.

Finlay, B.B., Frost, L.S., and Paranchych, W. (1986b) Origin of transfer of IncF plasmids and nucleotide sequences of the type II *oriT*, *traM*, *traY* alleles from ColB4-K98 and the type IV *traY* allele from R100-1. *J. Bacteriol*. **168**: 132-139.

Finlay, B.B., Frost, L.S., and Paranchych, W. (1986c) Nucleotide sequence of the *traYALE* region from IncFV plasmid pED208. *J. Bacteriol*. **168**: 990-998.

Finnegan, D.J., and Willetts, N.S. (1971) The site of action of the F transfer inhibitor. *Mol. Gen. Genet.* **111**: 256-264.

Fowler, T., Taylor, L., and Thompson, R. (1983) The control region of the F plasmid transfer operon: DNA sequence of the *traJ* and *traY* genes and characterization of the *traYZ* promoter. *Gene* **26**: 79-89.

Frost, L.S., Finlay, B.B., Opgenorth, A., Paranchych, W., and Lee, J.S. (1985) Characterization and sequence analysis of pilin from F-like plasmids. *J. Bacteriol.* **164**: 1238-1247.

Garner , M. M., and Revzin, A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**: 3047-3060.

Hewick, R. M., Hunkapiller, M. W., Hood, L. E., Dreyer, W. J. (1981) A gas-liquid solid phase peptide and protein sequenator. *J. Biol. Chem.* **256**: 7990-7997.

Hunkapiller, M., Lujan, E., Ostrander, F., and Hood, L. (1983) Isolation of microgram quantities of proteins from polyacrylamide gels for amino acid sequence analysis. *Meth. Enzymol.* **91**: 227-236.

Inamoto, S., Yoshioka, Y., and Ohtsubo, E. (1988) Identification and characterization of the products from the *traJ* and *traY* genes of plasmid R100. *J. Bacteriol.* **170**: 2749-2757.

Inamoto, S., and Ohtsubo, E. (1990) Specific binding of the TraY protein to *oriT* and the promoter region for the *traY* Gene of plasmid R100. *J. Biol. Chem.* **265**: 6461-6466.

Ippen-Ihler, K., and Minkley, E.G. (1986) The Conjugation system of F, the fertility factor of *Escherichia coli. Ann. Rev. Genet.* **20**: 593-624.

Kleckner, N. (1981) Transposable elements in prokaryotes. *Ann. Rev. Genetics* **15**: 341-404.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680-685.

Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951) Protein measurement with the Folin phenol reagent. *J. Biol. Chem.* **193**: 265-275.

Lugtenberg, B., Meyers, J., Peters, R., v. d. Hoeck, P., and v. Alphens, L. (1975) Electrophoretic resolution of the major outer membrane protein of *Escherichia coli* K-12 into four bands. *FEBS Letters* **58**: 254-258.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Maxam, A.M., and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Meth. Enzymol.*. **65**: 499-580.

Miller, J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Ohlsson, H., and Edlund, T. (1986) Sequence-specific interactions of nuclear factors with the insulin gene enhancer. *Cell* **45**: 35-44.

Paranchych, W., Finlay, B.B., and Frost, L.S. (1986) Studies on the regulation of IncF plasmid transfer operon expression. In *Banbury Report 24: Antibiotic Resistance Genes: Ecology, Transfer, and Expression.* Levy, S.B., and Novick, R.P. (eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 117-129.

Paranchych, W., and Frost, L.S. (1988) The physiology and biochemistry of pili. *Adv. Microbiol. Physiol.* **29**: 53-114.

Ronson, C.W., Nixon, B.T., and Ausubel, F.M. (1987) Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. *Cell* **49**: 579-581.

Russel, M., and Model, P. (1984) Replacement of the *fip* gene of *Escherichia coli* by an inactive gene cloned on a plasmid. *J. Bacteriol.* **159**: 1034-1039.

Tabor, S., and Richardson, C. C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**: 1074-1078.

Takeda, Y., Ohlendorf, D.H., Anderson, W.F., and Matthews, B.W. (1983) DNA-binding proteins. *Science* **221**: 1020-1026.

Thompson, R., and Taylor, L. (1982) Promoter mapping and DNA sequencing of the F plasmid transfer genes *traM* and *traJ*. *Mol. Gen. Genet.* **188**: 513-518.

Towbin, H., Staehelin, T., Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**: 4350-4354.

Traxler, B.A., and Minkley, E.G. (1988) Evidence that DNA helicase I and *oriT* site specific nicking are both functions of the F TraI protein. *J. Mol. Biol.* **204**: 205-209.

Vieira, J., and Messing, J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**: 259-268.

Willetts, N.S., and Skurray, R. (1987) Structure and function of the F factor and mechanism of conjugation. In Escherischia coli *and* Salmonella typhimurium*:*

Cellular and Molecular Biology. Ingraham, J., *et al* (eds.), Amer. Soc. for Microbiology, Washington, D.C., pp. 1110-1133.

Yamato, I., Anraku, Y., Hirosawa, K. (1975) Cytoplasmic membrane vescicles of *Escherichia coli*. *J. Biochem*. **77**: 705-718.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

CHAPTER III

The TraM Protein of the Conjugative Plasmid pED208 Recognizes a Sequence with Dyad Symmetry

A. Introduction

The mechanism of initiation of DNA transfer during bacterial conjugation is beginning to be elucidated. The *oriT* region is the site of action of the nicking complex TraYI (Inamoto *et al.*, 1991; Matson and Morton, 1991; Reygers *et al.*, 1991), the host factor IHF (Inamoto *et al.*, 1990; Tsai *et al.*, 1990; Dempsey and Fee, 1990) and the putative signal protein TraM (Abo *et al*, 1991; Schwab *et al.*, 1991; Chapter II). Genetic studies showed that TraM is essential for transfer but dispensable for nicking at *oriT* or pilus synthesis (Everett and Willetts, 1980). Therefore, TraM has been postulated to signal that a stable mating pair has formed and DNA transfer to the recipient cell can start (Kingsman and Willetts, 1978).

TraM from the conjugative plasmid pED208, the derepressed form of the IncFV plasmid F_0*lac*, binds to three sites in the *oriT* region characterized by 16 *Hin*fI-like sites of sequence GANTC, spaced 11-12 base pair (bp) apart or multiples thereof (Fig. III. 1). All of these sites except one (nucleotides 280-284) were protected from DNase I digestion as shown in Chapter II; it was proposed that the *Hin*fI-like sites were either the actual TraM binding sites or an integral part of the binding sites. To elucidate the role of these sites in the interaction of TraM with the *oriT* region, a series of oligonucleotides containing the *Hin*fI sites GATTC and/or

GAATC in different combinations and with different spacing were synthesized, and their interactions with TraM were examined.

B. Experimental Procedures

1. Oligonucleotide synthesis

Synthetic oligonucleotides corresponding to the sequences shown in Table III. 1 were synthesized on an Applied Biosystem DNA synthesizer (Mod. 391EP) by Pat Murray in the DNA synthesis facility (Dr. Ken Roy's laboratory) in the Department of Microbiology and purified by denaturing polyacrylamide gel electrophoresis and electroelution. The 22mer and 23mer oligonucleotides in Table III. 1 were designed such that after hybridization there would be a two-nucleotide single-stranded tail of CG on each 5' end.

2. Preparation of labeled oligonucleotides

The oligonucleotides were labeled at the 5' end as follows: 0.3 OD_{260} of DNA was resuspended into 5 µl of 10x polynucleotide kinase buffer (500 mM Tris-HCl pH 7.6, 100 mM $MgCl_2$, 50 mM dithiothreitol, 1 mM spermidine, 1 mM disodium EDTA) plus 10 µl of $\gamma[^{32}P]$-ATP (Dupont, Lachine, Quebec, Canada) and 35 µl of glass distilled water. One unit of T4 polynucleotide kinase was added and the mixture was incubated at 37°C for 45 min; the enzyme was inactivated by incubation at 65°C for 10 min and the labeled DNA was separated from the unincorporated label by means of reverse chromatography using Sep-pak C_{18} cartridges (Waters Associates, Milford, MA). The cartridges were conditioned with 10 ml of acetonitrile, and equilibrated with glass distilled water before loading the kinase mixture which had been diluted to 1 ml with 10 mM Tris-HCl, 1 mM disodium EDTA pH 8.0 (TE buffer). The cartridges were then washed with 25 mM ammonium bicarbonate, followed by 25 mM ammonium bicarbonate in 5% acetonitrile, followed by 5% acetonitrile. The labeled DNA was eluted with four 200 µl aliquots of 30% acetonitrile. The samples containing the largest amount of radioactivity were dried, resuspended in 20 µl each of TE buffer and pooled. The final volume of the labeled DNA was brought to 100 µl.

3. Preparation of double-stranded oligonucleotides

Preparation of the double-stranded species was performed as follows: equal amounts of the complementary synthetic oligonucleotides were mixed together in DNase I footprinting buffer (25 mM HEPES pH 7.8, 50 mM KCl, 0.5 mM DTT, 0.05 mM EDTA, 5% glycerol, 0.5 mM phenylmethylsulphonyl fluoride (PMSF)) and incubated for 30 minutes at 95°C. The incubator was then shut off and the mixture was allowed to cool to room temperature. The double-stranded DNA was then used for the DNA electrophoretic mobility retardation assays either as a labeled substrate for TraM binding or as a non-labeled competitor. The concentrations of oligonucleotides used were determined using the Beer and Lambert law: $Abs_{260}=\varepsilon lc$, where ε is the molar extinction coefficent calculated from the nucleotide composition and equal to 10768 and 10813 $cm^{-1}M^{-1}$ for the 22- and 23-nucleotide long primers, respectively; l is the pathlength of the cell and is equal to 1 cm, and c is the molar concentration.

4. Electrophoretic mobility retardation assay

The DNA-protein complexes were analysed using the method described in Chapter II. The ^{32}P radioactively labeled DNA was used for the electrophoretic mobility retardation assays and the DNA-protein complexes were separated on 5% or 12% non-denaturing polyacrylamide gels. Electrophoresis was performed at 20 mA at room temperature. The dried gels were exposed to X-ray film with intensifying screens (Eastman Kodak Co, Rochester, NY) at -70°C.

5. Circular dichroism

Circular dichroism (CD) measurements were carried out by Kim Oikawa in the laboratory of Dr. C. Kay on a Jasco J-720 spectropolarimeter (Jasco Inc., Easton, MD) interfaced to an Epson Equity 386/25 and controlled by Jasco software according to the protocol in Chapter V. Protein concentrations ranged from 0.5 to 0.7 mg ml^{-1}, DNA concentrations ranged from 0.5 to 1.0 mg ml^{-1} and the complexes ranged from 0.6 to 0.8 mg ml^{-1}. The mean residue weight for TraM was 116.024 (determined from the amino acid sequence reported in Chapter II); for DNA it was 304.545, and appropriate values for the complexes were applied. The unit for molar ellipticity was degree centimeter squared per decimole.

C. Results

1. Choice of synthetic oligonucleotides

In Chapter II, the regions protected by TraM from DNase I digestion were called footprints I, II and III. In this Chapter the nomenclature introduced by Abo *et al.* (1991) was adopted, and the sites were renamed *sbmA*, *sbmB* and *sbmC*, respectively, where sbm stands for site of specific binding of the TraM protein. The linear map of the *oriT* region of pED208 and the sequence of the TraM binding sites are shown in Figure III. 1.

The first sequence to be recognized by TraM, *sbmA*, contains 4 *Hin*fI sites (recognition sequence GANTC) spaced, in pairs, 12 bases apart, whereas the *sbmB* sequence contains 6 *Hin*fI sites plus one site with a non conserved base (Fig. III. 1, underlined with a broken line), which are spaced 11 bases apart with the exception of three sites that are spaced 12 and 7 bases apart (Fig. III. 1). The *sbmC* site contains three *HinfI* sites and two sites with one non conserved base which are spaced 11 bases apart except for the second and third sites and the fourth and the fifth sites which are 22 and 9 bases apart, respectively. Note that the second last *Hin*fI-like site (GACTA) in *sbmC* was not identified in the previous Chapter.

The sequences GATTC and GAATC are present in these regions 11 times in 16 sites, making them ideal candidates for the binding sites of TraM.

In addition to the binding studies with TraM and natural *oriT* DNA sequences presented in Chapter II, synthetic oligonucleotides containing one or two *Hin*fI sites spaced 11 or 12 bases apart were synthesized for use in binding studies and their sequences are shown in Table III. 1. The sequence between the *Hin*fI-like sites in the *oriT* region (Fig. III. 1) appeared to be non-specific, thus the sequence between the *Hin*fI sites in the synthetic DNA was arbitrarily chosen to be ATCGAT, corresponding to the restriction site *Cla*I. Because of the self-complementary nature of the sequence of one of the 22 nucleotide long oligonucleotides (template **T-A**,

Figure III. 1

Linear map of the *oriT* region of the conjugative plasmid pED208 and nucleotide sequence of the TraM binding sites. Top: a linear map of the *Eco*RI/*Pst*I fragment containing the *Eco*RI/*Sau*3A fragment from the *oriT* region. The thick bar represents the *Sau*3A/*Pst*I fragment from the multiple cloning site of the vector pUC18 (Yanisch-Perron *et al.*, 1985). The regions protected by TraM from DNase I digestion are shown by the striped boxes and are named according to the nomenclature introduced by Abo *et al.* (1991). The arrowheads and the dotted box within *sbmC* represent the hypersensitive sites and a region of partial protection respectively as discussed in Chapter II. Bottom: the sequence of the three TraM footprints is reported and is based on the data presented in Chapter II. The *Hin*fI sites are underlined with a continuous line whereas the sequences with one non-identical base are underlined with a broken line. The distance between neighboring sites is shown.

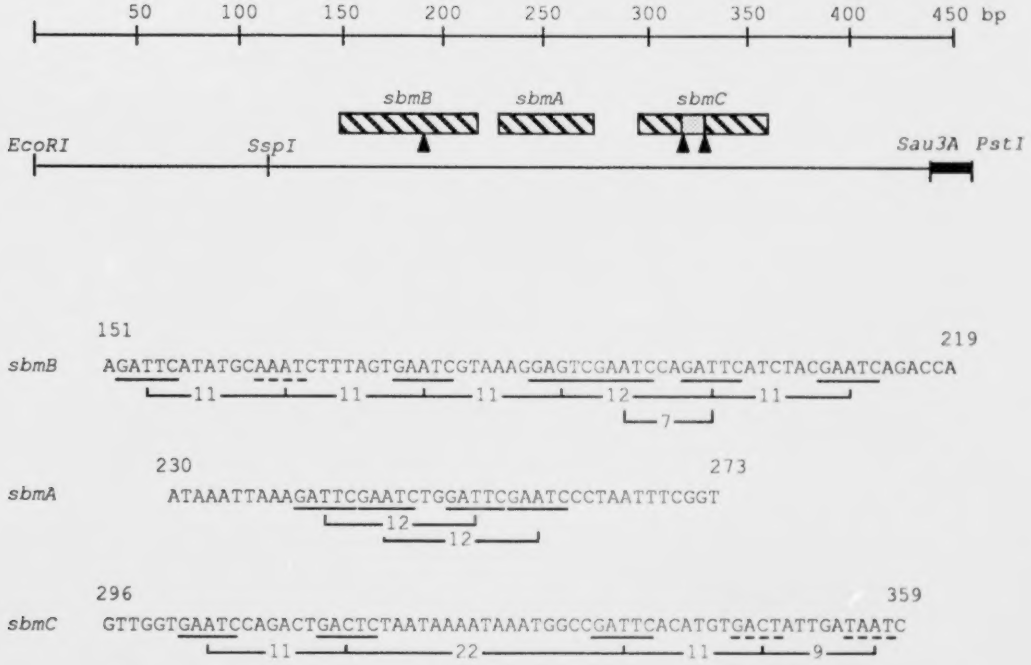

450 bp
sbmB
sbmA
sbmC
EcoRI
SspI
Sau3A
PstI
sbmB
AGATTCATATGCAAATCTTTAGTGAATCGTAAAGGAGTCGAATCCAGATTCATCTACGAATCAGACCA
sbmA
ATAAATTAAAGATTCGAATCTGGATTCGAATCCCTAATTTCGGT
sbmC
GTTGGTGAATCCAGACTGACTCTAATAAAATAAATGGCCGATTCACATGTGACTATTGATAATC

Table III. 1

Synthetic oligonucleotide sequences

Oligo*	Sequence (5'-3')
A	GAT**GAATC**GTCG
T	CGAC**GATTC**ATC
T-A	CGAC**GATTC**ATCGAT**GAATC**GT
T-T	CGAC**GATTC**ATCGAT**GATTC**GT
A-A	CGAC**GAATC**ATCGAT**GAATC**GT
T-T-**T**	CGAC**GATTC**TATCGAT**GATTC**GT
A-A-**A**	CGAC**GAATC**ATCGATA**GAATC**GT

* For easy identification, the oligonucleotides were named according to the nature of the nucleotide at position 3 in the *Hinf*I site, i.e., primers containing the sequence GAATC were named **A**, whereas primers containing the sequence GATTC were named **T**.

The double-stranded forms of the oligonucleotides are as follows:

T/A duplex:	$^{5'}$GAT**GAATC**GTCG$^{3'}$ $^{3'}$CTA**CTTAG**CAGC$^{5'}$
T-A duplex:	$^{5'}$CGAC**GATTC**ATCGAT**GAATC**GT$^{3'}$ $^{3'}$TG**CTAAG**TAGCTA**CTTAG**CAGC$^{5'}$
T-T/A-A duplex:	$^{5'}$CGAC**GATTC**ATCGAT**GATTC**GT$^{3'}$ $^{3'}$TG**CTAAG**TAGCTA**CTAAG**CAGC$^{5'}$
T-T-**T/A**-A-**A** duplex:	$^{5'}$CGAC**GATTC**TATCGAT**GATTC**GT$^{3'}$ $^{3'}$TG**CTAAG**ATAGCTA**CTTAG**CAGC$^{5'}$

Table III. 1), only three primers were synthesized. The 23-nucleotide long oligonucleotides were obtained by addition of either one extra T residue between the 5' end *Hin*fI site and the *Cla*I site in the **T**-T-**T** oligonucleotide, or one extra A residue between the *Cla*I site and the 3' end *Hin*fI site in the **A**-A-**A** oligonucleotide (underlined in Table III. 1).

2. DNA electrophoretic mobility retardation assays

The labeled oligonucleotides were incubated with purified TraM protein in order to determine the minimal binding sequence recognized by the protein. Sucrose gradient studies (Chapter V) had indicated a molecular weight for TraM of approximately 60,000, suggesting that the protein is a tetramer. Thus a molecular weight of 58,000 (4x14,500) was adopted and the amounts of TraM used in the following experiments were calculated according to this value. When purified TraM was incubated with primers **A** and **T**, either as single- or double-stranded DNA, no binding was detected at high TraM concentrations (data not shown). TraM did not bind to the single-stranded form of the oligonucleotides carrying two *Hin*fI sites except for oligonucleotide **T**-**T**; a constant amount of the oligonucleotide showed lower electrophoretic mobility when incubated with TraM over a wide range of protein concentrations (Fig. III. 2). Electrophoretic mobility retardation assays of the double-stranded form of the longer primers showed binding of TraM to all of the oligonucleotides at a molar ratio of protein to DNA of approximately 2:1 (Fig. III. 3). Thus TraM recognizes two *Hin*fI sites spaced 11 or 12 bases apart.

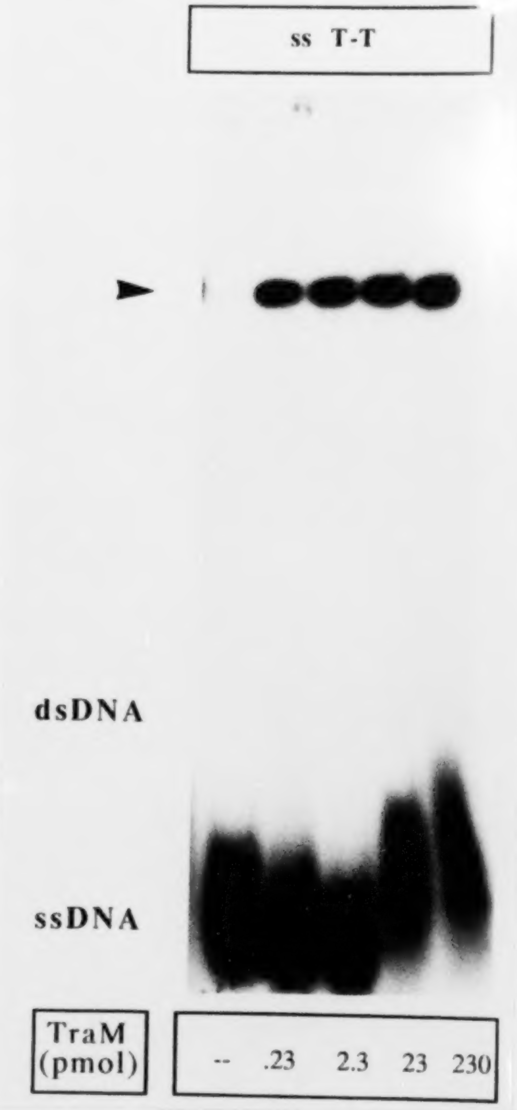

ss T-T
dsDNA
ssDNA
TraM
(pmol)
--
.23
2.3
23
230

Figure III. 2

Binding of TraM to the single-stranded oligonucleotide T-T. Two pmol of ^{32}P-labeled DNA were incubated with increasing amounts of TraM as shown in the box at the bottom of the Figure. The mobility of single- and double-stranded DNA, as well as that of bound DNA is indicated.

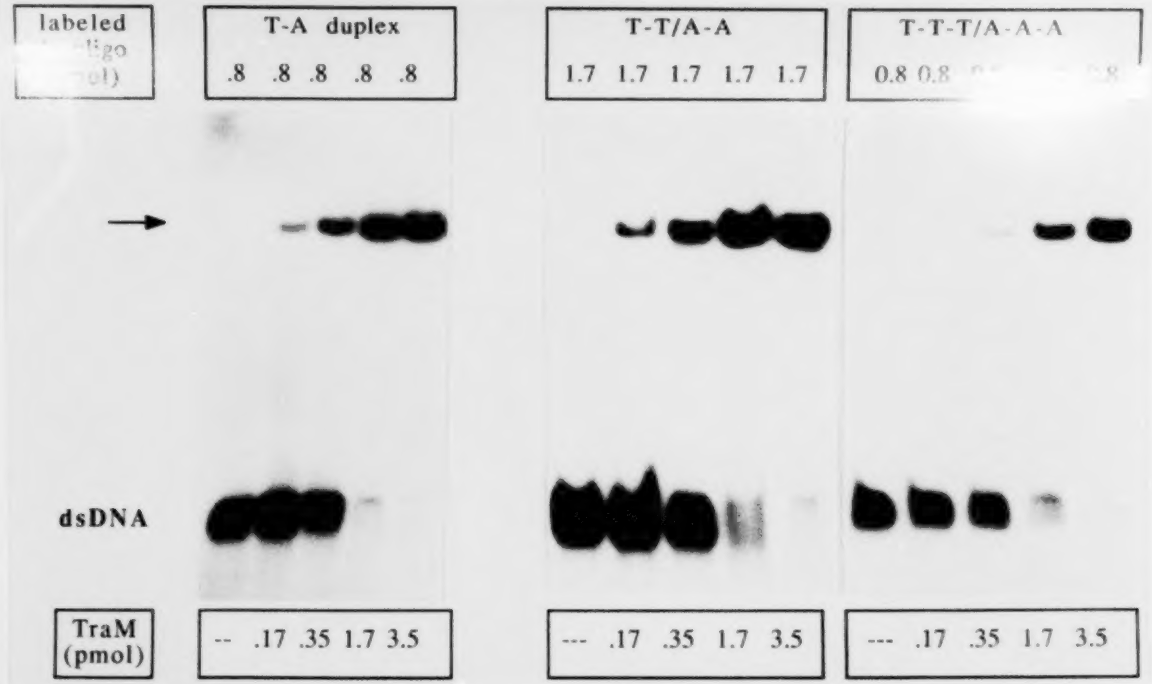

labeled
T-A duplex
.8 .8 .8 .8 .8
T-T/A-A
1.7 1.7 1.7 1.7 1.7
T-T-T/A-A-A
0.8 0.8
dsDNA
TraM
(pmol)
-- .17 .35 1.7 3.5
--- .17 .35 1.7 3.5
--- .17 .35 1.7 3.5

Figure III. 3

Binding of TraM to the double-stranded oligonucleotide pairs **T-A** duplex, **T-T/A-A** duplex and **T-T-T/A-A-A** duplex. The amounts of radioactively labeled DNA and the corresponding amount of TraM used in the incubation are indicated above and below the autoradiograms, respectively. The mobilities of double-stranded DNA and of DNA-protein complexes are indicated.

3. Competition assays

In order to elucidate which of the oligonucleotides was specific for the DNA-binding domain of TraM, electrophoretic mobility retardation assays of labeled oligonucleotides were performed in the presence of the unlabeled oligonucleotides.

Competition assays were performed using the oligonucleotides in double-stranded form as shown in Figure III. 4. The molar ratios of protein to labeled synthetic DNA were approximately 4:1 (**T-A** duplex and **T-T-T/A-A-A**) and 2:1 (**T-T/A-A**). Oligonucleotide **T-T/A-A** duplex inhibited binding of TraM to template **T-T-T/A-A-A** at very low concentrations (Fig. III. 4, right panel, lane 1) and the same template inhibited binding of TraM to template **T-A** duplex at high concentration (Fig. III. 4, left panel, lane 4). Even at high concentration, oligonucleotide **T-T-T/A-A-A** was an inefficient inhibitor of TraM binding to the oligonucleotide **T-T/A-A** (Fig. III. 4, central panel, lane 4).

The oligonucleotides were also used in competition assays with DNA fragments from the *oriT* region of pED208.

The radioactively labeled *Eco*RI/*Pst*I fragment from pLDL101 (Chapter II, page 49), containing the three TraM binding sites (Fig. III. 1), was incubated with TraM in the absence or presence of increasing amounts of the single-stranded oligonucleotide **T-T**. Complete inhibition of binding of the protein to the *oriT* region at a molar ratio of protein to synthetic DNA of approximately 1:4 was observed, suggesting specific interaction of TraM with the synthetic single-stranded DNA (data not shown).

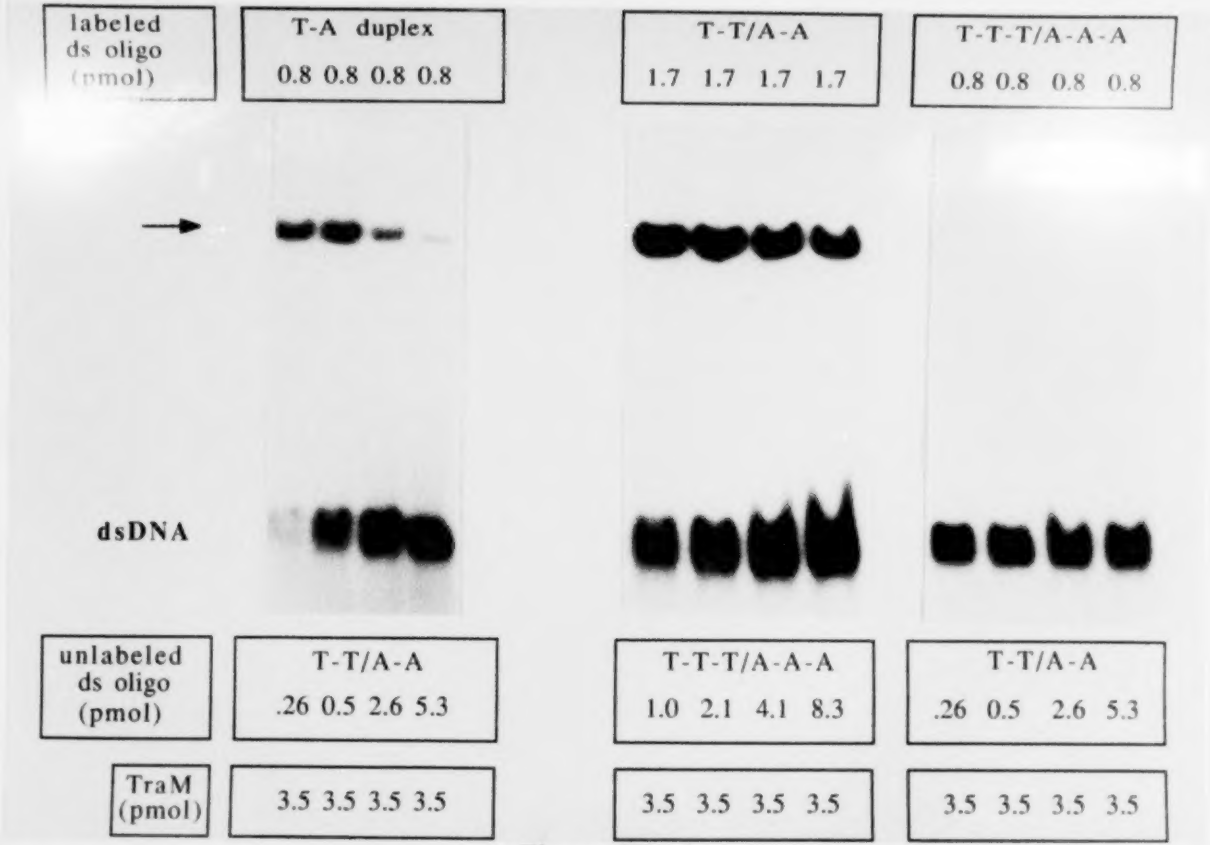

labeled ds oligo (pmol)
T-A duplex
0.8 0.8 0.8 0.8
T-T/A-A
1.7 1.7 1.7 1.7
T-T-T/A-A-A
0.8 0.8 0.8 0.8
dsDNA
unlabeled ds oligo (pmol)
T-T/A-A
.26 0.5 2.6 5.3
T-T-T/A-A-A
1.0 2.1 4.1 8.3
T-T/A-A
.26 0.5 2.6 5.3
TraM (pmol)
3.5 3.5 3.5 3.5
3.5 3.5 3.5 3.5
3.5 3.5 3.5 3.5

Figure III. 4

Competition of TraM binding to the double-stranded oligonucleotides. The amounts of radioactively labeled DNA used in the incubation reactions are indicated on the top. The amounts of TraM and non-labeled competitor DNA used in each reaction are indicated at the bottom. The mobilities of the double-stranded DNA and of DNA-protein complexes are shown.

The labeled DNA fragment containing the *oriT* region was incubated with quantities of TraM that saturated the binding sites in the *oriT* region (Fig. III. 5, lane 3) in the presence of double-stranded oligonucleotides. Both oligonucleotides **T-A** duplex and **T-T/A-A** duplex were able to inhibit binding of TraM to the labeled *oriT* region (Fig. III. 5, lanes 6 and 9), where oligonucleotide **T-T/A-A** was capable of inhibiting binding of TraM to the *oriT* region at a lower concentration compared to oligonucleotide **T-A** duplex (compare Fig. III. 5, lanes 5 and 7). While oligonucleotide **T-T-T/A-A-A** duplex could inhibit binding at lower concentrations of TraM (Fig. III. 5, lane 10) and it was partially effective at high protein concentrations (Fig. III. 5, lane 12). This suggested that TraM has high affinity for double-stranded oligonucleotides carrying two *HinfI* sites 11 bases apart. A summary of the results of the binding studies is presented in Table III. 2.

4. Circular dichroism

Circular dichroism (CD) has been used as a sensitive probe for determining the secondary structure of macromolecules. It is based on the ability of asymmetric molecules to rotate the plane of polarized light. Two types of information can be derived from the ultra-violet CD spectra of proteins: a) the secondary structural content of the molecule and b) the conformation of the environment of surrounding aromatic groups. The effect of secondary structures is more prominent in the far UV-CD spectra (190-250 nm), whereas the aromatic residues primarily affect the near UV-CD spectra (250-320 nm) For a more detailed characterization of TraM see Chapter V. CD can be used to characterize DNA-protein interactions; in particular, CD is sensitive to conformational changes upon complex formation. Circular dichroism studies of TraM incubated in the presence of either single-stranded

Figure III. 5

Inhibition of TraM binding to the *oriT* region by the synthetic oligonucleotides. One femtomole of the labeled *Eco*RI/*Pst*I DNA fragment containing the *oriT* region was incubated with the amounts of TraM and non-labeled double-stranded oligonucleotides indicated at the bottom of the Figure. The mobility of the free DNA fragment is shown.

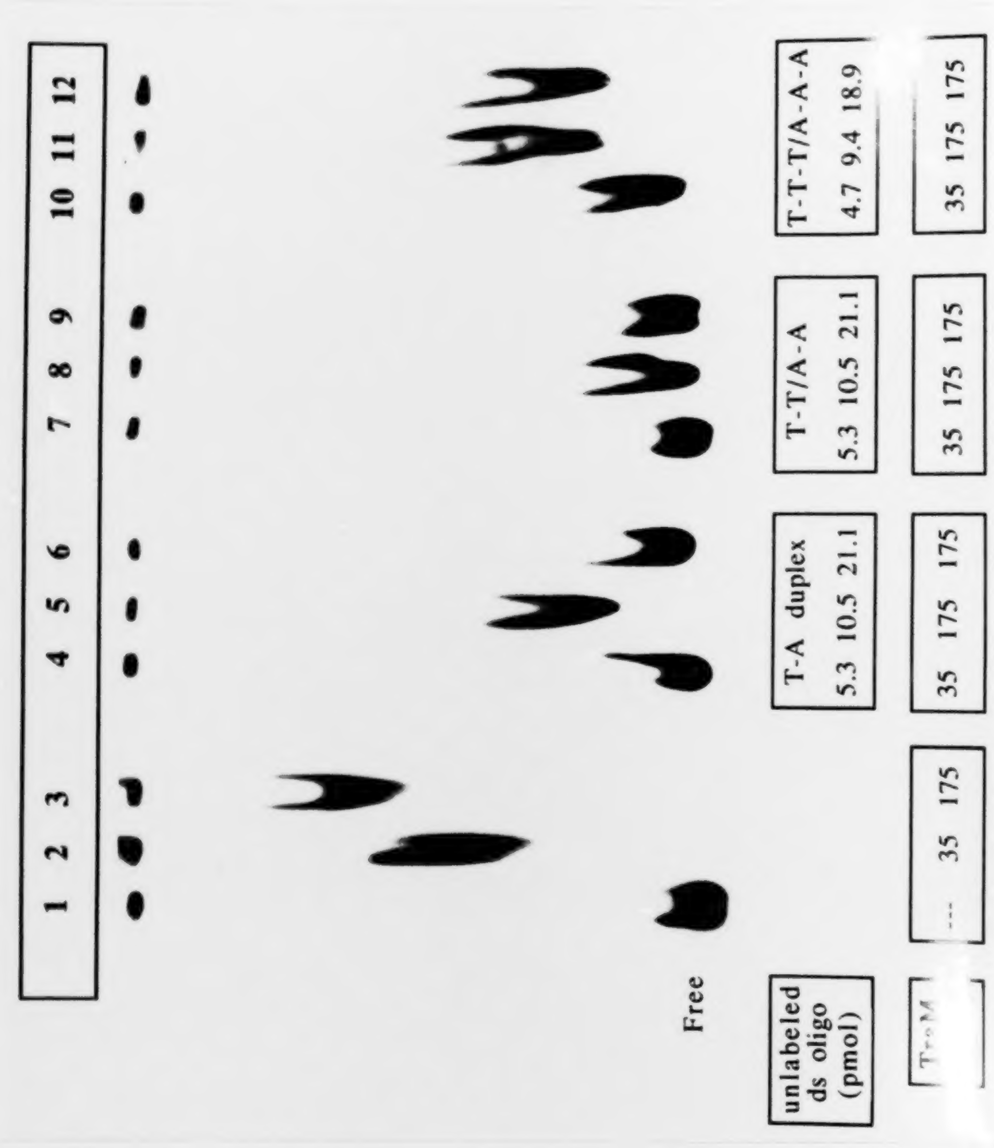

1 2 3 4 5 6 7 8 9 10 11 12
Free
unlabeled ds oligo (pmol)
T-A duplex
5.3 10.5 21.1
T-T/A-A
5.3 10.5 21.1
T-T-T/A-A-A
4.7 9.4 18.9
--- 35 175
35 175 175
35 175 175
35 175 175

Table III. 2

Summary of the interactions of TraM with the different oligonucleotides

Oligonucleotide	Sequence (5'-3')	TraM binding
A	GAT**GAATC**GTCG	--
T	CGAC**GATTC**ATC	--
T-A	CGACGATTCATCGAT**GAATC**GT	--
T-T	CGAC**GATTC**ATCGAT**GATTC**GT	+
A-A	CGAC**GAATC**ATCGAT**GAATC**GT	--
T-T-T	CGACGATTCTATCGAT**GATTC**GT	--
A-A-A	CGAC**GAATC**ATCGATA**GAATC**GT	--
A/T duplex		--
T-A duplex		+++
T-T/A-A duplex		+++
T-T-T/A-A-A duplex		++

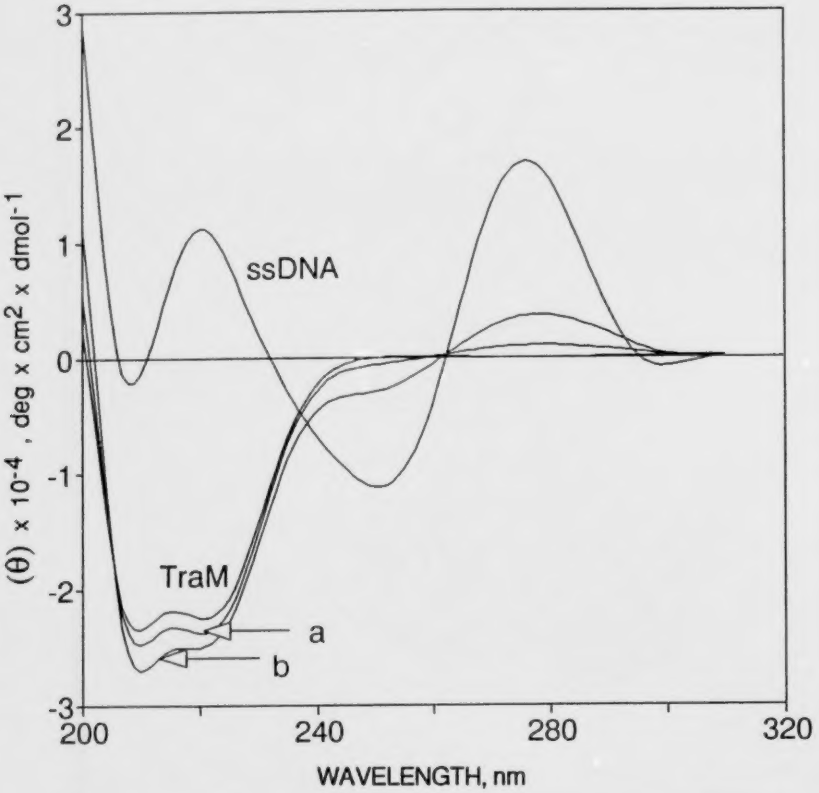
oligonucleotide T-T or double-stranded oligonucleotide T-T/A-A were performed as described in the Experimental Procedures.

The molar ratio of protein to single-stranded template was 1:0.96, 1:1.92, 1:2.66 and 1:3.84. Figure III. 6 shows the spectra of only two of these experiments as well as the spectra of the single components. The CD difference spectra (Δ-spectra) (Fig. III. 6, bottom graph) were determined by subtracting the spectra obtained by summation of the values of the ellipticity of each component of the complex at any given wavelength (calculated spectra) from the experimentally observed spectra. The CD spectra of the mixture of TraM incubated with the double-stranded oligonucleotide T-T/A-A were determined at a molar ratio of protein to DNA of 1:0.52, 1:1.03, 1:1.56 and 1:2.07. Figure III. 7 shows the spectra of TraM and double-stranded DNA, and the spectra of two of the protein-DNA complexes, as well as their corresponding Δ-spectra (Fig. III. 7, bottom graph).

The α-helical band at 220 nm of the protein-DNA mixtures showed increased negative ellipticity compared to the ellipticity of the protein alone (-22,470). This was contrary to the expected result, since the calculated spectra for non-interacting components, obtained by simple graphical summation of the ellipticities of each component, would have been higher (smaller absolute value) due to the positive ellipticity of the DNA in this range of wavelengths (DNA is a right-handed helix) (Fig. III. 6 and 7) as shown in Table III. 3. The negative ellipticity of the DNA-protein complexes increased linearly as the concentration of oligonucleotides increased, reaching saturation at a molar ratio of protein to DNA of approximately

Figure III. 6

CD spectra of the interaction of TraM with the single-stranded oligonucleotide **T-T**. The molar ratios of protein to DNA were (a) 1: 0.96 and (b) 1: 3.84 in both plots. Top: CD spectra of the protein incubated with the single-stranded oligonucleotide. The CD spectra of the protein alone as well as of the DNA are indicated. Bottom: difference spectrum (Δ-spectrum) of the protein incubated with the single-stranded oligonucleotide **T-T**.


3
2
1
0
-1
-2
-3
(θ) x 10-4 , deg x cm2 x dmol-1
ssDNA
TraM
a
b
200
240
280
320
WAVELENGTH, nm

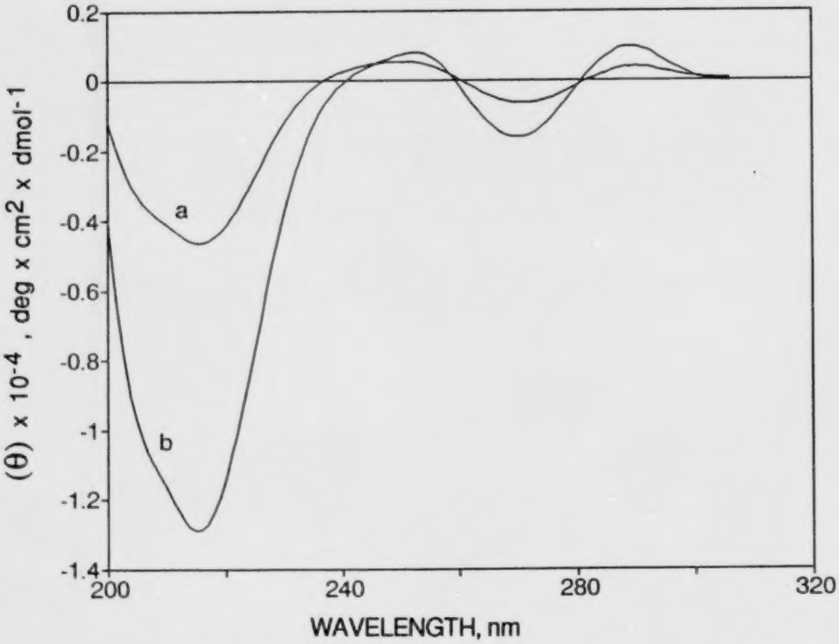
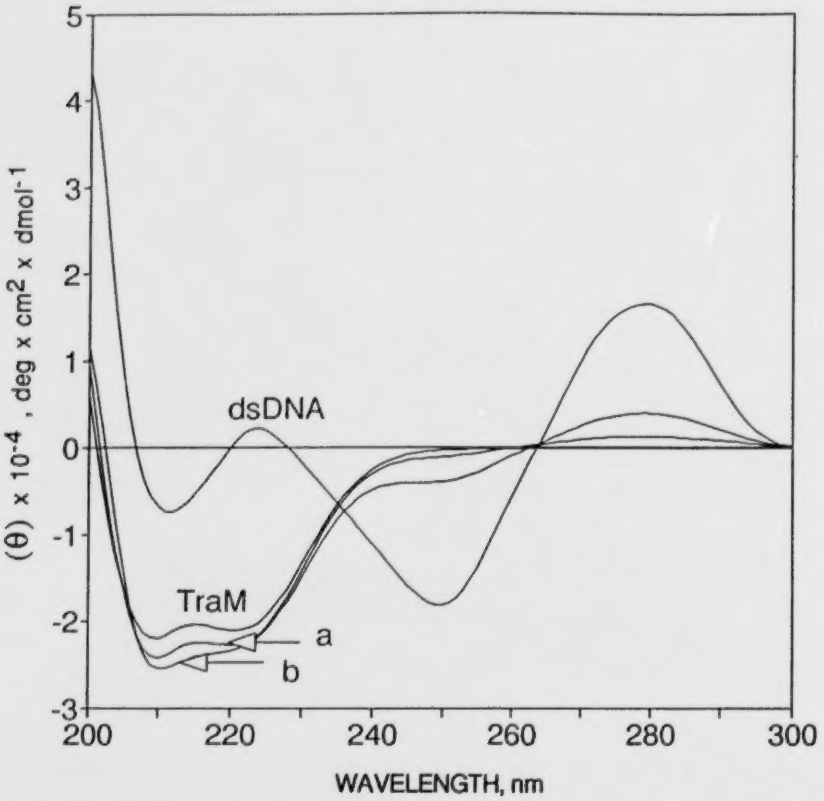

0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
(θ) x 10-4 , deg x cm2 x dmol-1
a
b
200
240
280
320
WAVELENGTH, nm

Figure III. 7

CD spectra of the interaction of TraM with the double-stranded oligonucleotide **T-T/A-A**. The molar ratios of protein to DNA were (a) 1: 0.52 and (b) 1: 2.07 in both plots. Top: CD spectra of the protein incubated with the double-stranded oligonucleotide. The CD spectra of the protein alone as well as of the double-stranded DNA are indicated. Bottom: difference spectrum (Δ-spectrum) of the protein incubated with the double-stranded oligonucleotide **T-T/A-A**.


5
4
3
2
1
0
-1
-2
-3
(θ) x 10-4 , deg x cm2 x dmol-1
dsDNA
TraM
a
b
200
220
240
260
280
300
WAVELENGTH, nm

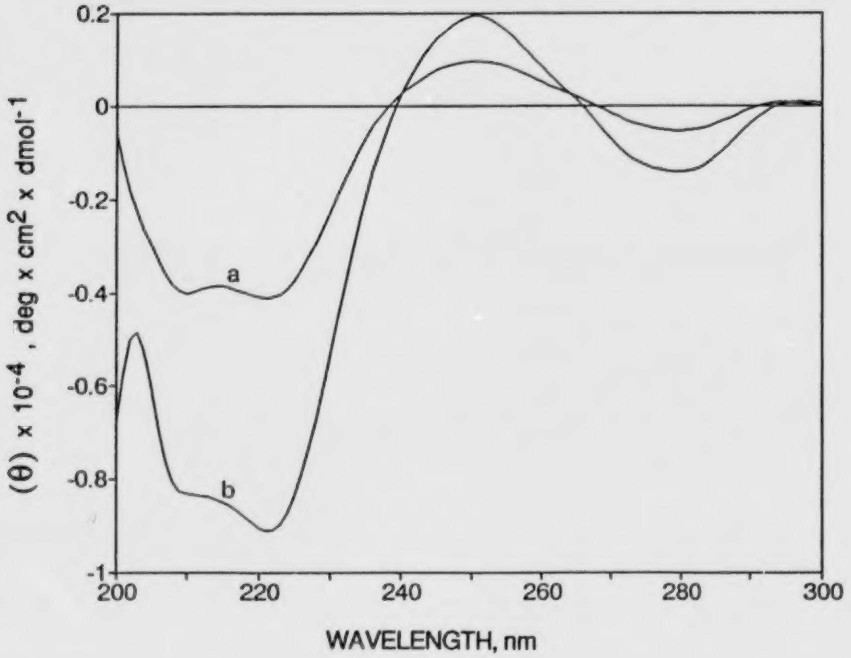

0.2
0
-0.2
-0.4
-0.6
-0.8
-1
(θ) x 10-4 , deg x cm2 x dmol-1
a
b
200
220
240
260
280
300
WAVELENGTH, nm

Table III. 3

Ellipticity values of DNA-protein complexes at 220 nm

	(θ) observed	(θ) calculated	Δ-(θ)
TraM	-22,470		
ssDNA*	+11,220		
dsDNA*	-65		
TraM: ssDNA			
1:0.96	-23,780	-19,140	-4,640
1:1.92	-24,610	-16,410	-8,200
1:2.66	-24,770	-13,460	-11,310
1:3.84	-25,070	-12,200	-12,870
TraM: dsDNA			
1:0.52	-22,750	-18,820	-3,930
1:1.03	-23,890	-17,060	-6,830
1:1.56	-23,930	-15,550	-8,380
1:2.07	-23,420	-14,520	-8,900

*ssDNA indicates the single-stranded oligonucleotide **T-T**, whereas dsDNA indicates the double-stranded oligonucleotide **T-T/A-A**.

1:2.07. These results suggest the presence of a conformational transition upon interaction of the protein and the DNA.

D. Discussion

Seven single-stranded oligonucleotides and the four double-stranded molecules derived from their annealing reactions were analyzed for their ability to be recognized by TraM of pED208.

The protein favours the double-stranded oligonucleotide that had two copies of the *HinfI* site 11 or 12 bases apart over the oligonucleotides with one *HinfI* site. This is consistent with the results from DNaseI protection of the *oriT* region shown in Chapter II. TraM protects all the *HinfI*-like sites in the region except one that does not have a "partner" at a correct distance. Furthermore, the spacing between the two *HinfI* sites is important, since the protein prefers the oligonucleotides that have the *HinfI* sites 11 bases apart. The extra base in oligonucleotides **T**-T-**T/A**-A-**A** adds a rotation of 34°-36° to the two sequences relative to their positions in the DNA helix, which places the *HinfI* sites out of phase resulting in a lower affinity of TraM for the oligonucleotide **T**-T-**T/A**-A-**A**. Such results contrast with what is seen in the *oriT* region, where TraM binds first to the *sbmA* site which has two pairs of *HinfI* sites 12 bp apart. It is possible that the high affinity of TraM for *sbmA* is the result of the unique arrangement of the two binding sites within *sbmA* which are overlapping (Fig. III. 1). Another possibility is that the sequences around *sbmA* alter the conformation of the DNA or contribute to the interaction of TraM with this site. However, protein flexibility rather than conformational changes in the structure of the DNA has been proposed to explain binding of proteins to oligonucleotides with dyad symmetry that have one extra base in between the symmetrical sequences (Weiss *et al.*, 1990; Sadler *et al.*, 1983).

TraM also shows some affinity for the single-stranded form of **T-T**. The reasons for such interactions are not known. It is possible that the oligonucleotide may form double-stranded structures, although bands of electrophoretic mobility corresponding to double-stranded DNA were not detected.

The competition assays showed that both the single- and the double-stranded form of the oligonucleotides containing two GATTC sequences 11 bases apart, can effectively compete with the binding of TraM to the *oriT* region, suggesting that these oligonucleotides may represent the optimal TraM binding site.

Circular dichroism performed on the complexes between TraM and the oligonucleotides showed a conformational transition. It is very difficult to assign the changes observed in the CD spectra to either one of the two components of the complex. In many different DNA-binding protein systems, such as the yeast transcriptional activator GCN4 (Weiss *et al.*, 1990; Talanian *et al.*, 1990), the DNA binding domain of the "leucine zipper" family (O'Neil *et al.*, 1990) and the heterodimer of the proto-oncogene products Fos and Jun (Patel *et al.*, 1990), transition to higher negative ellipticity at 220 nm of protein-DNA complexes has been attributed to an increased α-helical content of the polypeptide upon binding the DNA; according to this interpretation, TraM may have changed its α-helical content in the complex. To test this hypothesis, the CD spectra of TraM in the presence of the strong helix-inducer trifluoroethanol (TFE) (50% v/v final concentration) were performed. TFE is a highly polar molecule which perturbs the aqueous environment surrounding polypeptides. As a result, the intrachain interactions of the residues are stabilized and the proteins are believed to assume their maximum α-helical conformation. The CD spectra of TraM in the presence of TFE shows increased negative ellipticity comparable to that observed in the protein-DNA complex (data not shown). However, an increase in negative ellipticity may also be explained by changes in the quaternary structures of proteins, for example slight changes in the relative orientation of two helices within the same polypeptide (Strynadka and James, 1989) or rotation of side-chains may result in different ellipticity (Yang, 1965). Furthermore, changes in the conformation of the DNA molecule cannot be excluded, since the Δ-spectra of the protein-DNA complexes in the near UV, which is sensitive to the bases of the DNA and, to a lesser extent, to the aromatic residues of the protein, shows appreciable variation.

The results obtained in this work suggest a mechanism for TraM binding to the *oriT* region of pED208. TraM may bind to seven sites in the *oriT* region as represented schematically in Figure III. 8, interacting with the GA-TC nucleotides of the binding sites. The reason for the apparent higher affinity of TraM for the oligonucleotides **T-T/A-A** over **T-A** duplex is presently unclear. One possible explanation is that maximum interactions between the protein and the synthetic DNA are reached when position 3 of the oligonucleotide is occupied by a T residue. X-ray crystallography of the TraM protein alone or in complex with the oligonucleotide may clarify the paradox between the synthetic DNA data presented in this Chapter and the DNase I protection analysis of the *oriT* region previously shown.

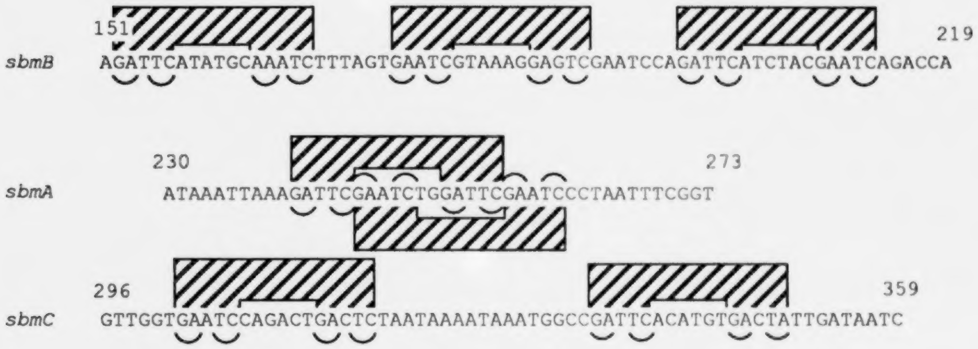

151
219
sbmB
AGATTCATATGCAAATCTTTAGTGAATCGTAAAGGAGTCGAATCCAGATTCATCTACGAATCAGACCA
230
273
sbmA
ATAAATTAAAGATTCGAATCTGGATTCGAATCCCTAATTTCGGT
296
359
sbmC
GTTGGTGAATCCAGACTGACTCTAATAAAATAAATGGCCGATTCACATGTGACTATTGATAATC

Figure III. 8

Model of TraM binding to the *oriT* region of pED208. The TraM protein is represented by the striped box and the conserved nucleotides are circled.

E. Bibliography

Abo, T., Inamoto, S. and Ohtsubo, E. (1991) Specific DNA Binding of the TraM Protein to the *oriT* Region of Plasmid R100. *J. Bacteriol.* **173**: 6347-6354.

Everett, R. and Willetts, N. (1980) Characterisation of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136**: 129-150.

Inamoto, S., Abo, T. and Ohtsubo, E. (1990) Binding sites of Integration Host Factor in *oriT* of plasmid R100. *J. Gen. Appl. Microbiol.* **36**: 287-293.

Inamoto, S., Yoshioka, Y. and Ohtsubo, E. (1991) Site- and strand-specific nicking *in vitro* at *oriT* by the TraY-TraI endonuclease of plasmid R100. *J. Biol. Chem.* **266**: 10086-10092.

Kingsman, A. and Willetts, N. (1978) The requirements for conjugal DNA synthesis in the donor strain during F*lac* transfer. *J. Mol. Biol.* **122**: 287-300.

Matson, S.W. and Morton, B. S. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J. Biol. Chem.* **266**: 16232-16237.

O'Neil, K.T., Hoess, R. H. and DeGrado, W. F. (1990) Design of DNA-Binding Peptides Based on the Leucine Zipper Motif. *Science* **249**: 774-778.

Patel, L., Abate, C. and Curran, T. (1990) Altered protein conformation on DNA binding by Fos and Jun. *Nature* **347**: 572-575.

Reygers, U., Wessel, R., Müller, H. and Hoffmann-Berling, H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J.* **10**: 2689-2694.

Sadler, J.R., Sasmor, H. and Betz, J. L. (1983) A perfectly symmetric *lac* operator binds the *lac* repressor very tightly. *Proc. Natl. Acad. Sci. U.S.A.* **80**: 6785-6789.

Schwab, M., Gruber, H. and Hogenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5**: 439-446.

Strynadka, N.C.J. and James, M.N.G. (1989) Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* **58**: 951-998.

Talanian, R.V., McKnight, C. J. and Kim, P. S. (1990) Sequence-Specific DNA Binding by a Short Peptide Dimer. *Science* **249**: 769-771.

Tsai, M-M., Fu, Y. H. F. and Deonier, R. C. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J. Bacteriol.* **172**: 4603-4609.

Weiss, M.A., Ellenberger, T., Wobbe, C. R., Lee, J. P. Harrison, S. C. and Struhl, K. (1990) Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. *Nature* **347**: 575-578.

Yang, J.T., (1965) On the phenomenological treatments of optical rotary dispersion of polypeptides and proteins. *Proc. Natl. Acad. Sci. U.S.A.* **53:** 438-445.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

CHAPTER IV

The TraM Protein of the Conjugative Plasmid F Binds to the Origin of Transfer of the F and ColE1 Plasmids[1]

A. Introduction

In F and F-like plasmids there are more than 30 genes involved in conjugation and they are organized in three major transcriptional units: *traM*, *traJ* and the *traYI* operon (as reviewed in Ippen-Ihler, 1989; Willetts and Skurray, 1987). The origin of transfer or *oriT* region of the F plasmid contains the nick site, which has been remapped by Reygers *et al.* (1991), and is the site of action of the plasmid-encoded TraI (Traxler and Minkley, 1988), TraY (Lahue and Matson, 1990) and TraM proteins as well as Integration Host Factor, IHF (Tsai *et al.*, 1990). The *traM* gene product is thought to be involved in signalling the formation of a stable mating pair and in the initiation of DNA transfer (Kingsman and Willetts, 1978). The ColE1 plasmid is a well known non-conjugative plasmid, whose origin of replication has been used in molecular biology to create the most commonly employed plasmid vectors such as pBR322 (Bolivar *et al.*, 1977). This plasmid is mobilized efficiently by conjugative plasmids belonging to the IncF, IncI and IncP groups (Willetts and Wilkins, 1984). The mechanism by which this process occurs is not known, but probably involves some of the *tra* functions.

---

[1] A version of this chapter has been submitted for publication. Di Laurenzio, L., Frost, L. S., and Paranchych W. Mol. Microbiology. 1992.

It was shown in Chapter II that the *traM* gene product of the F-like plasmid pED208 is a DNA binding protein: the interaction of F TraM protein with the *oriT* regions of F and ColE1 plasmids is investigated in this Chapter.

B. Experimental Procedures

1. Bacterial strains, media and plasmids

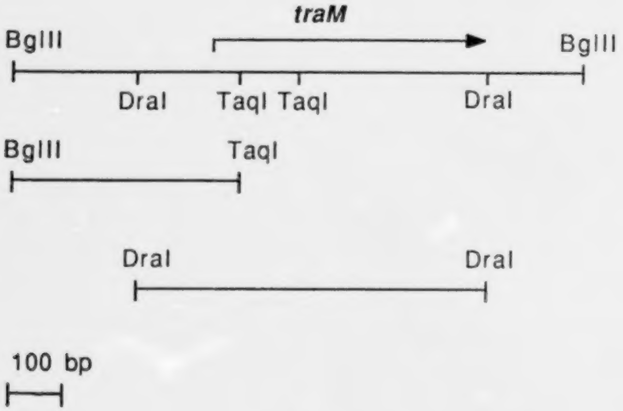
Bacterial strains used in this work are shown in Table IV.1. *E. coli* cells were grown in Luria-Bertani medium supplemented accordingly with ampicillin to a final concentration of 25 μg ml^{-1}.

2. Cloning and over-expression of the *traM* gene

The *DraI* fragment containing the *traM* gene (Fig. IV. 1) was cloned into the *SmaI* site within the multiple cloning site of pUC18 and the resulting plasmid was called pLDLF7. The *Eco*RI/*Hin*dIII fragment from pLDLF7 containing the *traM* gene was cloned into the expression vector pT7-4 (Tabor and Richardson, 1985) and named pLDLF007. In order to over-express the *traM* gene product, *E. coli* BL21(DE3) cells were transformed with pLDLF007. This *E. coli* strain carries the T7 polymerase gene under the control of the p_{tac} promoter (Studier and Moffat, 1986). Over-expression was achieved by addition of isopropyl-β-thiogalactopyranoside (IPTG, Sigma Chemicals, St Louis, MO) (final concentration 2 mM) to 2 l of overnight grown cultures. Cells were ready for F-TraM protein purification after 90 min incubation at 37°C.

Table IV. 1.

Bacterial strain	genotype	Source/Reference
Escherichia coli		
JC3272	*lac*$^{-}$Sm^{s}Spc^{r}T6^{r}P1^{r}	Achtman *et al.*, 1971
M176	JCFLO/JC3272	*ibidem*
BL21	F^{-}*hsd*S*gal*	Studier,W. F. and Moffat, B. A., 1986
DH5α	*end*A*hsd*R17(r_k-r_m+)*sup*E*thi*-1*rec*A *gyr*A*rel*AΔlacU(φ80d*lac*ZΔM15)	Sambrook *et al.*, 1989


traM
BglII
BglII
DraI
TaqI TaqI
DraI
BglII
TaqI
DraI
DraI
100 bp

Plasmid	Vector
pNY300	pUC18
pLDLF13	pUC18
pLDLF7	pUC18
pLDLF007	pT7-4

Figure IV. 1

Linear map of the clones used in this Chapter. pNY300 is described in Frost *et al.* (1989); pLDLF13 was created by cloning the *Bgl*II/*Taq*I fragment into the *Bam*HI/*Acc*I sites of pUC18; pLDLF7 was created by cloning the purified *Dra*I fragment into the *Sma*I site of pUC18 and pLDLF007 was created by subcloning the *Eco*RI/*Hin*dIII fragment from pLDLF7 into the *Eco*RI/*Hin*dIII sites of pT7-4.

3. TraM purification

Purification of the over-expressed protein was carried out as previously described (Chapter II) with the following modifications. Fractions from anion exchange chromatography (FPLC MonoQ HR10/10) containing the TraM protein were pooled and dialysed against 50 mM Tris-HCl (pH 8), 0.1 mM ethylenediamine tetraacetic acid disodium salt (EDTA), 1 mM dithiothreitol (DTT) (TED buffer) and concentrated 10 times using Microsep microconcentrators with a molecular cut-off of 10 kDa (Filtron, Northborough, MA). Two milliliter aliquots were applied to a five milliliter bed volume column of Affi-Gel Blue (BioRad, Richmond, CA). The column was washed with two bed volumes of TED buffer and stepwise elution was performed using 5 ml of each of the following buffers: 0.5 M NaCl/TED, 0.8 M NaCl/TED, 0.9 M NaCl/TED, 1.0 M NaCl/TED, 1.2 M NaCl/TED, 2.0 M NaCl/TED and 4.0 M NaCl/TED. The purified protein was recovered from the 1.2 and 2.0 M NaCl/TED fractions. After extensive dialysis against 0.1 M NaCl/TED, the sample was concentrated to 2 mg ml^{-1} and stored frozen at -20°C and withstood several freeze-thaw cycles. Automated protein sequence analysis of the purified protein was performed by Mike Carpenter (Department of Biochemistry) as previously described (Chapter II).

4. Antisera preparation

Linda Glasier and Neil Ellert in Dr. Paranchych's laboratory (Department of Microbiology) assisted with antisera preparation. A Flemish Giant crossed with a French Lopped-ear rabbit was injected intramuscularly in the gluteal and subscapular areas with a total of 200 μg of purified protein from SDS gels (Jacobs and Clad, 1986) dissolved in equal volumes of sterile saline solution and Freund's complete adjuvant (3 ml total). The rabbit was injected two more times after 3 and 6 weeks with the same amount of protein and Freund's incomplete adjuvant. The animal was bled 2 weeks after final injection.

5. Detection of the end point of the anti TraM antisera

The titer of the anti-TraM antibodies was determined using enzyme-linked immunosorbant assay (ELISA). The principles of this assay have been described by Voller *et al.* (1974). Purified TraM protein was dissolved in coating buffer (0.5 M $NaHCO_3$, 0.1% NaN_3, pH 9.6) at a final concentration of 2 $\mu g\ ml^{-1}$. Two hundred microliter volumes of antigen in the coating buffer were placed in the wells of microtiter plates and incubated for 16 hr at 20°C in a moist atmosphere, the antigen saturated wells were washed three times for 3 min each in phosphate buffered saline containing 0.05% (v/v) Tween 20 (PBS-Tween). The wells were then filled with various dilutions of rabbit anti-TraM antiserum diluted in PBS-Tween containing 1% (w/v) BSA. Incubation with antiserum was for 2 hr at 20°C. The resulting antigen/antibody complexes were then incubated with alkaline phosphatase-conjugated goat-anti-rabbit antiserum (Boehringer-Mannheim, Laval, Quebec, Canada) and used at a 1:2000 dilution in PBS-Tween, the alkaline phosphatase substrate, p-nitrophenylphosphate (Sigma Chemicals, St Louis, MO) was added as a 1 mg ml^{-1} solution in 10% diethanolamine. The end point of the antiserum was defined as the highest dilution of antiserum that would cause absorbance at 405 nm of 0.02 OD and it was determined with a Titertek multiscan ELISA plate reader after an appropriate time interval following addition of the substrate for the color reaction.

6. Characterization and localization of TraM within cells

Fractionation of cells into cytoplasmic, inner and outer membrane fractions was carried out as previously described (Chapter II). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), protein transfer to nitrocellulose paper and immunological detection were performed as previously described (Chapter II).

7. Recombinant DNA techniques

Restriction enzymes, T4 DNA ligase, the large fragment of DNA polymerase I (Klenow) and T4 polynucleotide kinase were purchased from Boehringer-Mannheim. The enzymes were used according to suppliers specifications and protocols. DNA transfer to nitrocellulose filters, DNA filter hybridization, and colony filter hybridization have been described previously (Frost *et al.*, 1985; Finlay *et al.*, 1983). The transformation of *E. coli* with plasmid DNA and the isolation of DNA fragments for cloning procedures were according to Maniatis *et al.* (1982).

8. Characterization of protein-DNA interactions

The plasmid pLDLF13 was digested with *Hin*dIII and *Kpn*I, or with *Eco*RI and *Pst*I in order to generate fragments that could be selectively labeled at one end by the "filling-in" reaction (Maniatis *et al.*, 1982). Plasmid pBR322 was digested with *Pvu*II and *Nde*I. The 5' end of the *Nde*I site was dephosphorylated with calf intestinal alkaline phosphatase in order to generate an appropriate site for T4 polynucleotide kinase labeling (Maniatis *et al.*, 1982). The products of digestions were separated on 1.6% agarose gels and the fragments were purified according to the procedure of Dretzen *et al.* (1981). Labeling of the 3'- or 5'-ends of the fragments was performed with α-[^{32}P]-dATP or γ-[^{32}P]-ATP (Dupont, Lachine, Quebec, Canada), respectively. DNA-protein complex electrophoretic mobility retardation assays and DNase I footprinting were performed as previously described (Chapter II).

9. Formation of double-stranded synthetic oligonucleotides

A pair of synthetic oligonucleotides of sequence $^{5'}$GGGTGTCGGGGCGCAG CCATGACCC$^{3'}$ and its complementary strand were synthesized in the Department of Microbiology at the University of Alberta. Equal amounts of the two synthetic oligonucleotides were mixed together in DNase I footprint buffer (25 mM HEPES pH 7.8, 50 mM KCl, 0.5 mM DTT, 0.05 mM EDTA, 5% glycerol, 0.5 mM phenylmethylsulphonyl fluoride (PMSF)) and incubated for 30 minutes at 95^{o}C. The incubator was then shut off and the mixture was allowed to cool to room temperature. The double-stranded DNA was then used for the DNA electrophoretic mobility retardation assays either as substrate for binding or as a competitor.

C. Results

1. TraM over-expression and purification

TraM protein was over-expressed in *E. coli* as described in Experimental Procedures. SDS-PAGE of crude extracts from overproducing cells labeled with ^{35}S-methionine showed three major bands: two of which corresponded to the unprocessed and processed forms of the β-lactamase marker from the vector, while the third band, with electrophoretic mobility corresponding to a molecular weight of 10,950, was tentatively identified as TraM (data not shown). The TraM band was electro-eluted from SDS-PAGE as previously described (Chapter II) and its amino terminus sequenced. The sequence of the first five amino acids was MAKVN which corresponded to the sequence determined from the DNA (Thompson and Taylor, 1982), suggesting that the over-expressed protein was indeed TraM. The protein was purified as described in Experimental Procedures. The final yield of purified protein was 5 mg per liter of cell culture (Fig. IV. 2, lane 5). Sucrose gradients showed a molecular weight for TraM of approximately 60,000 suggesting that TraM is a tetramer under these conditions (Chapter V).

2. Immunological characterization of TraM

The purified protein was used to raise polyclonal antibodies as described in the Experimental Procedures. The titer of the antisera was determined using the Enzyme-Linked Immunosorbant Assay (ELISA) and corresponded to an end point of approximately 10^{-6}. The location of TraM protein in cells carrying the F plasmid and/or the *traM* gene on a multicopy plasmid, was determined using immunoblots of cell fractions separated by gel electrophoresis and transferred to nitrocellulose paper. *E. coli* JC3272 and M176 (JCFLO/JC3272, F^+), alone or transformed with plasmid pLDLF7 (Fig. IV. 1) were fractionated into cytoplasm, inner and outer membranes as described in Chapter II. The proteins were separated on duplicate denaturing 15% polyacrylamide gels and one of them was stained with Coomassie blue (data not shown), while the other was transferred to nitrocellulose paper. The transferred proteins were incubated with anti-TraM antiserum diluted 100-fold (Fig. IV. 3). The TraM protein was visible predominantly in the cytoplasm of all strains and there was a nearly equivalent amount of TraM in the inner membrane fraction of the cells carrying pLDLF7, as determined by densitometric scanning of the autoradiogram.

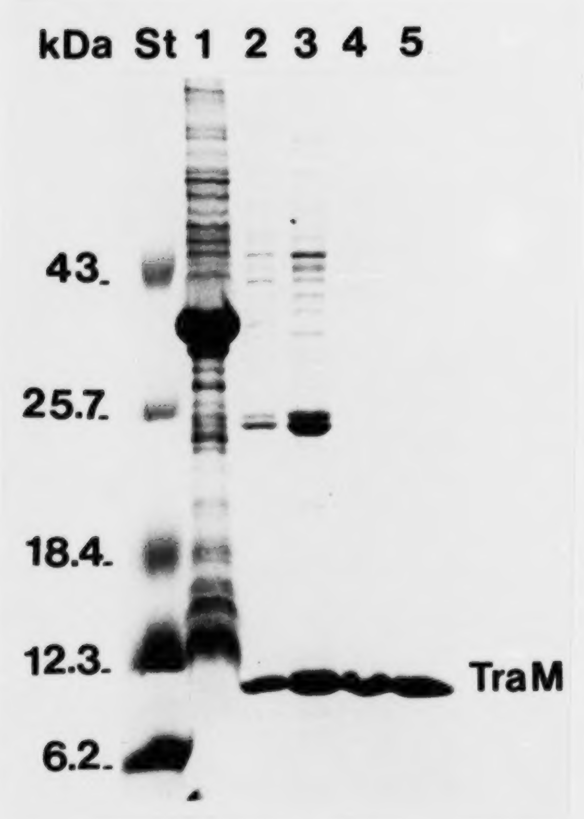


Figure IV. 2

Purification of the F TraM protein. Coomassie blue stain of SDS-polyacrylamide gel (15%) electrophoresis of fractions from the purification procedure. St., molecular weight standards; Lane 1, 30 µg of crude membrane preparation; Lane 2, 21 µg of soluble fraction; Lane 3, 30 µg of the supernatant of 50% ammonium sulfate fractionation; Lane 4, 18 µg of the peak eluted from the FPLC MonoQ; Lane 5, 10 µg of the 1.2-2.0 M NaCl fractions eluted from the Affi-Gel Blue.

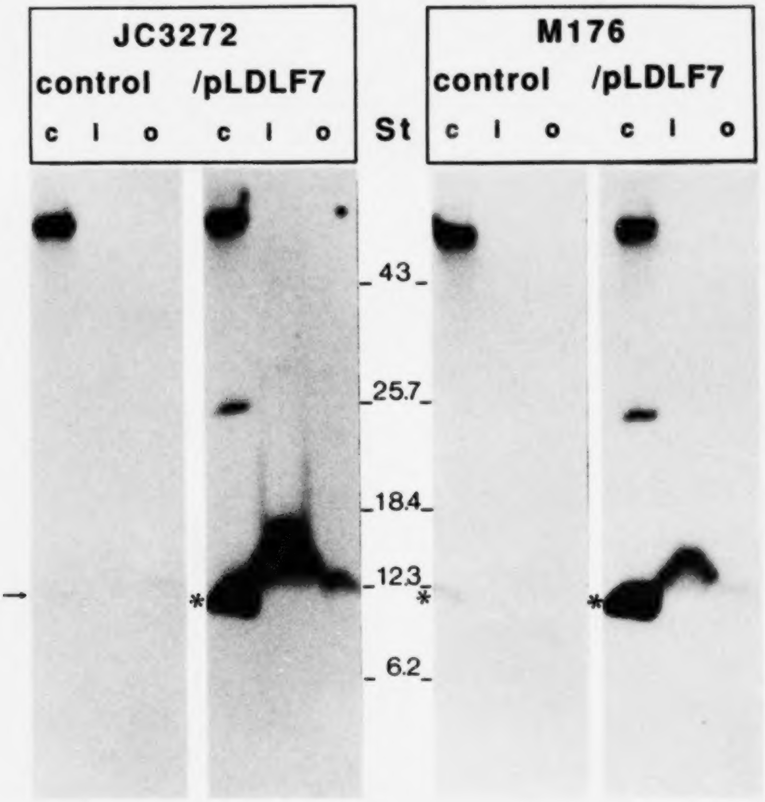

JC3272
control
/pLDLF7
c
i
o
c
i
o
St
M176
control
/pLDLF7
c
i
o
c
i
o
43
25.7
18.4
12.3
6.2
*
*
*

Figure IV. 3

Localization of the TraM protein. Immunoblot of cell fractions of cultures of JC3272, M176 alone or transformed with pLDLF7, separated into cytoplasmic (**c**), inner membrane (**i**), and outer membrane (**o**) fractions. Approximately 30 μg of protein from cytoplasmic and inner membrane fractions and 15 μg of outer membrane fraction were separated by SDS-PAG electrophoresis. St. Molecular weight standards expressed in kilodaltons. The asterisks indicate the TraM band, whereas the arrow indicates the mobility of the cross-reacting material discussed in the Results section.

The TraM band was barely visible in the M176 fractions, which contained a single copy of the *traM* gene, although the Coomassie blue stained duplicate showed similar total amounts of protein in each lane (data not shown). There was no apparent difference between the amount of TraM found in M176 and JC3272 transformed with the multicopy plasmid pLDLF7 suggesting that the presence of the F plasmid did not affect the cell localization of the TraM protein. Very weak cross-reactivity with a protein of molecular weight similar to TraM was detected in all the lanes of the control strain *E. coli* JC3272 after long exposures (Fig. IV. 3, indicated with an arrow). Densitometric scanning of the autoradiogram showed the intensity of the TraM band in the cytoplasmic fraction of M176, after correction for loading errors, to be twice the intensity of the background band in JC3272, whereas the bands in the inner and outer membrane fractions were of comparable intensity, suggesting that TraM is only present at detectable levels in the cytoplasm of strain M176. The band appearing at a molecular weight of 22,000 in the cytoplasmic fractions is a dimer of TraM, since it is present in immunoblots of purified antigen (see Fig V. 10) whereas the other bands of molecular weight above 43,000 are the products of non-specific cross-reactivity, since they were not present in immunoblots of purified antigen, but were present in the control strain JC3272. The nature of these bands is not known. The band in the outer membrane fraction of JC3272 transformed with pLDLF7, of molecular weight corresponding to TraM, is the result of overflow material from the adjacent well.

3. TraM-*oriT* complex formation

The purified protein was used in electrophoretic mobility retardation assays to demonstrate binding of F TraM to the *oriT* region of F. The *Eco*RI/*Pst*I fragment from plasmid pLDLF13 containing the *Bgl*II/*Taq*I fragment of the F *oriT* region (Fig. IV. 1) was selectively labeled at the 3' end of the *Eco*RI site with Klenow DNA polymerase. The labeled fragment was incubated with increasing amounts of purified F TraM protein. The complexes were separated on non-denaturing 5% polyacrylamide gels electrophoresed at room temperature (Fig. IV. 4). The TraM-*oriT* region complexes appeared as four discrete bands of decreasing mobility (Fig. IV. 4, panel F). The fourth band probably was not the result of a specific DNA-protein interaction, but rather the product of protein aggregation, since the protein tends to form aggregates at high concentration (note the concentration of TraM in lane 7 is 30 μM). Complex I was formed at a molar ratio of DNA to protein of 1: $3 \cdot 10^3$, complex II and complex III required two times and ten times more TraM, respectively. The binding constant for the first complex was estimated to be greater than 10^7 M^{-1}, which corresponds to the concentration of TraM in lane 2 of Figure IV. 4.

In 1984, Willetts and Wilkins identified a sequence within the *oriT* region of ColE1 that had high homology with the inverted repeats between nucleotides 240 and 275 in the F *oriT* region (Fig. IV. 7). A 220 bp *Pvu*II/*Nde*I fragment from pBR322 (Maniatis *et al.*, 1982), containing the region of homology (position 2299-2320 of the ColE1 sequence in Fig. IV. 7 A), was selectively labeled at the *Nde*I site and incubated with F TraM. The *oriT* region of the non-conjugative plasmid ColE1 showed electrophoretic mobility retardation of the band to two positions of lower mobility. The second band of very low mobility appeared at very high TraM concentrations, suggesting that protein aggregation might also have occurred in this sample (Fig. IV. 4, panel ColE1). We also tested the *oriT* region of the F-like

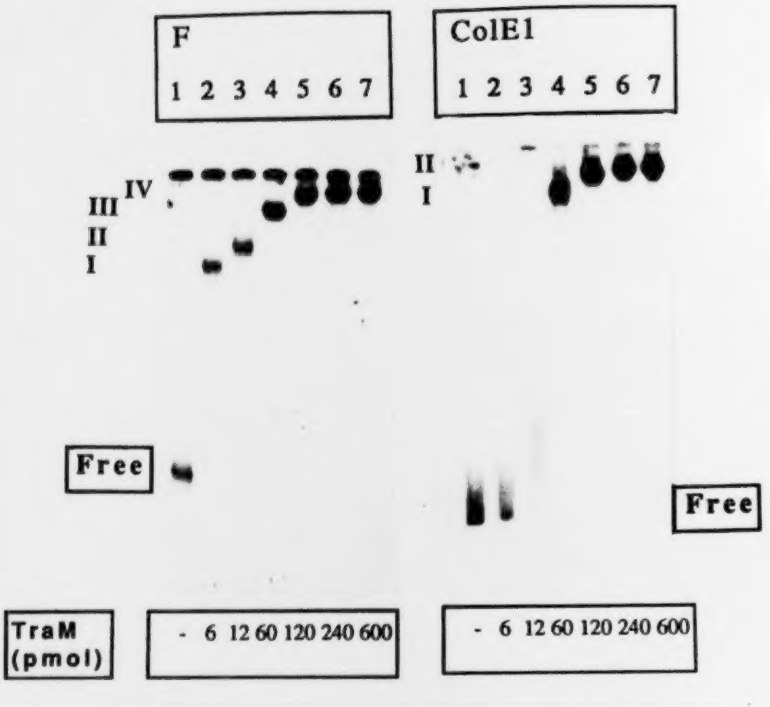

F
1 2 3 4 5 6 7
ColE1
1 2 3 4 5 6 7
IV
III
II
I
II
I
Free
Free
TraM (pmol)
- 6 12 60 120 240 600
- 6 12 60 120 240 600

Figure IV. 4

DNA electrophoretic mobility retardation assay of TraM-DNA complexes. DNA fragments of different origins were labeled at their 3' end as described in Experimental Procedures and incubated with increasing amounts of purified TraM (indicated in the box at the bottom of the figure) in the presence of 1 μg of poly(dI-dC)poly(dI-dC). The autoradiogram of the complexes separated on a 5% non-denaturing polyacrylamide gel is shown. **F.** 2.0 fmol of the 450 bp *Eco*RI/*Pst*I fragment of pLDLF13. **ColE1.** 2.0 fmol of the 229 bp *Pvu*II/*Nde*I fragment of pBR322 containing the *oriT* of ColE1. Free, mobility of free DNA fragment.

plasmid pED208, which did not show binding activity (data not shown).

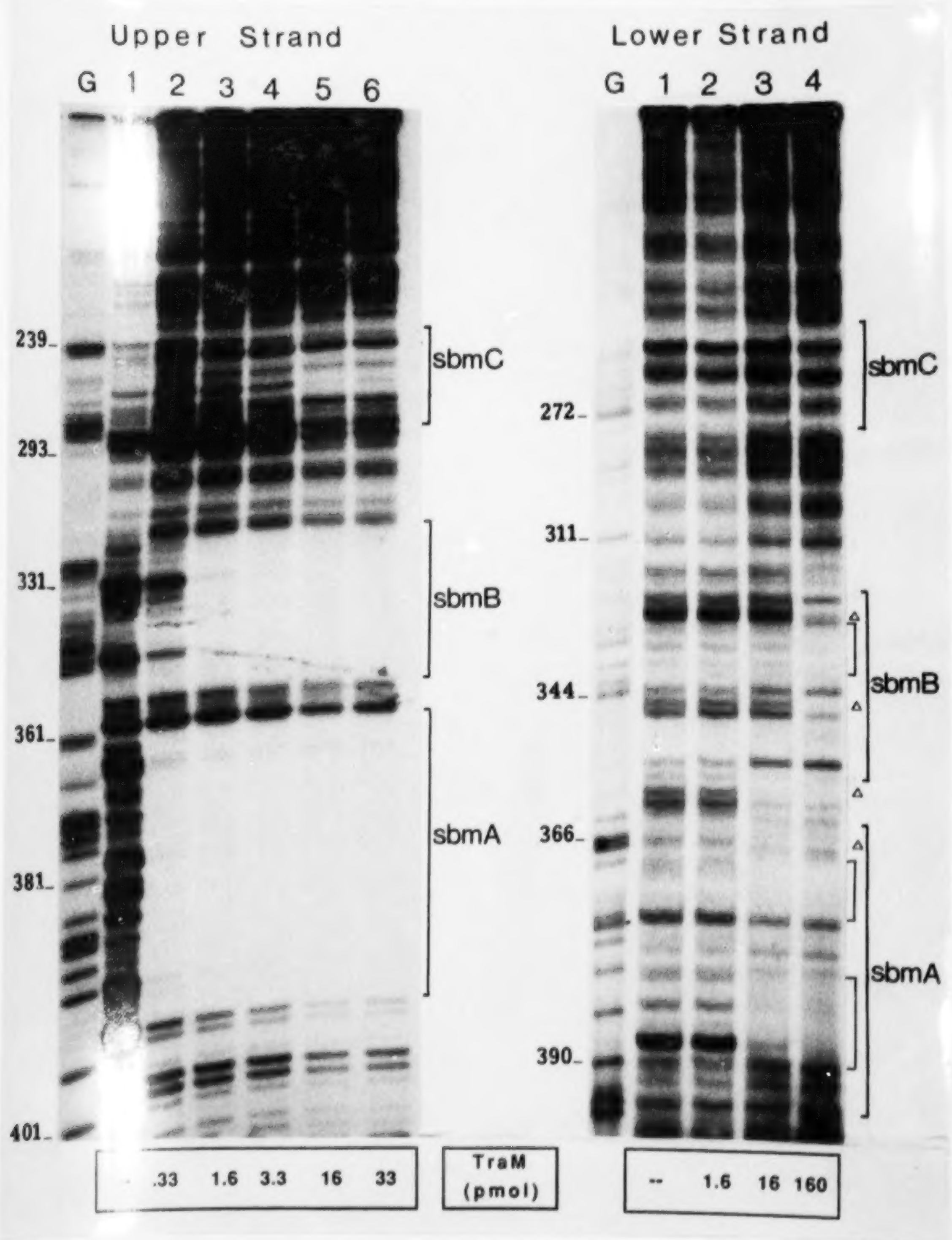
4. DNase I footprinting analysis of the F and ColE1 *oriT* regions with TraM protein

In order to localize the site(s) of interaction of the TraM protein with the *oriT* region, DNase I protection experiments were performed. The ^{32}P-labeled *Hin*dIII-*Kpn*I fragment from pLDLF13 was incubated in the presence of increasing amounts of purified TraM protein, briefly digested with DNase I, and the products of digestion were separated on 6% sequencing gels (Fig. IV. 5). Three main regions were protected: the first, in order of affinity for TraM, started at nucleotide 361 (*sbmA*, where sbm stands for specific binding site of the TraM protein according to the nomenclature introduced in Abo *et al.*, 1991) while the second was between nucleotide 322 and 349 (*sbmB*) and the third was between nucleotide 239 and 289 (*sbmC*), as the concentration of protein was increased. The first two footprints were localized in the TraM promoter region, *sbmA* covered the -10 region while *sbmB* covered the -35 region, whereas *sbmC* corresponded to a region of dyad symmetry initially described by Thompson and Taylor (1982). The three footprints appeared at protein to DNA molar ratios similar to the ones that caused formation of complexes I and II in the electrophoretic mobility retardation assays.

There was a different protection pattern between the upper and the lower strands. While the upper strand showed three footprints covering 34, 27 and 50 nucleotides (Fig. IV. 5, left panel), the lower strand was much more susceptible to DNase I digestion, even though five times more TraM was used in the incubation

Figure IV. 5

DNase I protection analysis of the *oriT* region in the presence of purified TraM. The *Hin*dIII/*Kpn*I (left panel) and *Eco*RI/*Pst*I (right panel) fragments from pLDLF13 containing the 422 bp fragment of F-*oriT* were labeled at the upper and lower strand respectively and incubated with purified TraM. Lane G (both panels), Maxam and Gilbert sequencing reaction. Lane 1 (both panels), DNase I digestion of 10.0 fmol of the DNA fragments. The footprinted regions are indicated with brackets, single base protection is shown by the open triangle.


Upper Strand
G 1 2 3 4 5 6
239
293
331
361
381
401
sbmC
sbmB
sbmA
- .33 1.6 3.3 16 33
TraM (pmol)
Lower Strand
G 1 2 3 4
272
311
344
366
390
sbmC
sbmB
sbmA
-- 1.6 16 160

mixture, and had very small protected regions ranging from a single base to a region of only a few bases (Fig. IV. 5, right panel), suggesting uneven interaction of the protein with the two DNA strands. The only protected regions of any length in the lower strand were the sequences complementary to the -35 region and to a lesser extent, the -10 region of the *traM* promoter.

We performed a similar experiment to investigate the interaction of the F TraM protein with the *oriT* region of the ColE1 plasmid. Unfortunately, the fragment was unstable and we were never able to obtain "footprints" of the site of interaction of TraM with the ColE1 *oriT* region. The reason for this instability is not known.

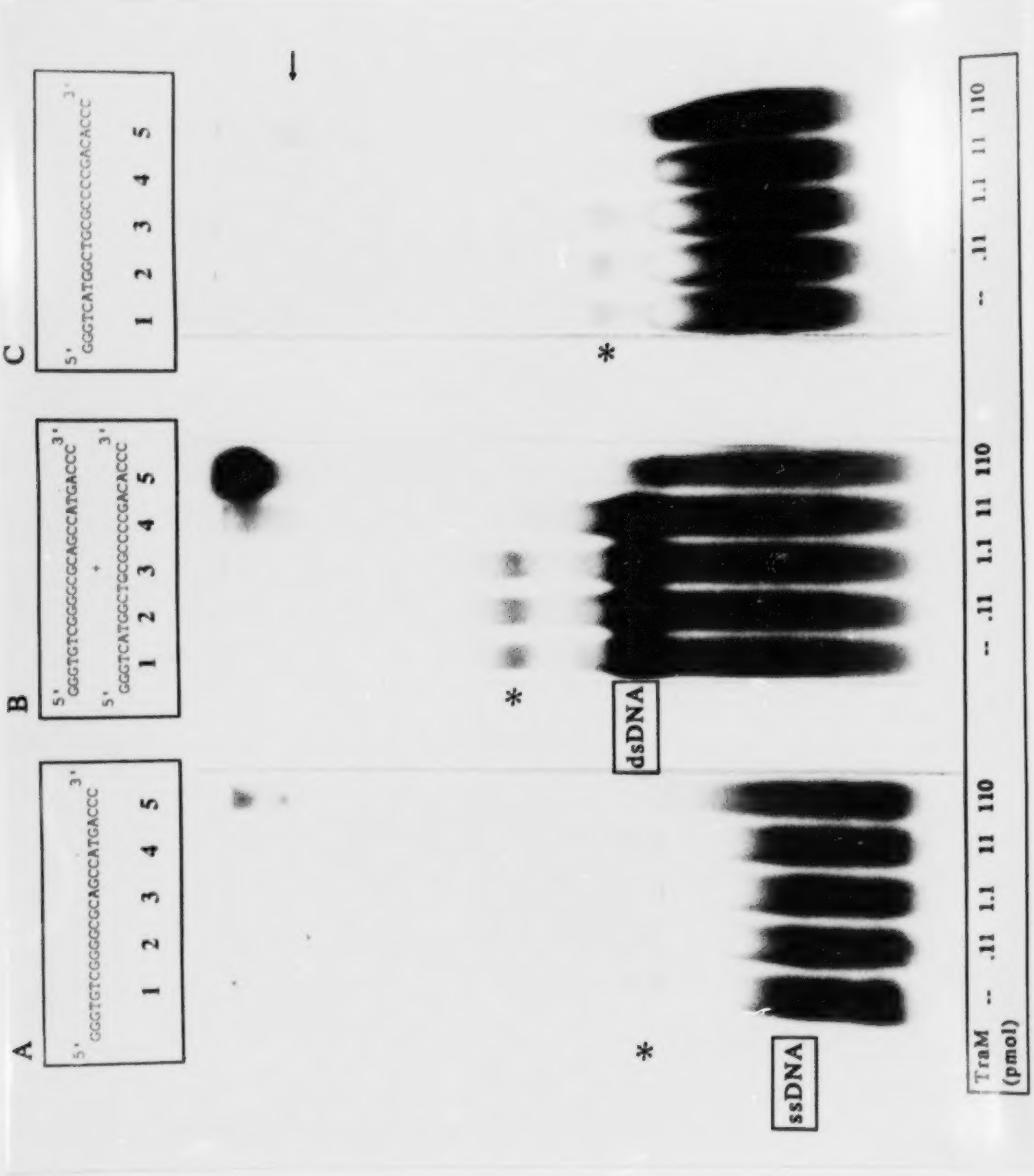
5. TraM binding to synthetic oligonucleotides

In order to investigate the interaction of TraM with the *oriT* region of ColE1, a pair of oligonucleotides was synthesized as described in Experimental Procedures. This pair of oligonucleotides corresponded to the ColE1 *oriT* region (position 2299-2320 of the ColE1 sequence in Fig. IV. 7 **A**) which is homologous to F *oriT* and its sequence was based on the pBR322 sequence published in Maniatis *et al.* (1982), which has a one base difference from the sequence in this region presented in Thompson *et al.* (1984).

Electrophoretic mobility retardation assays were performed on the labeled oligonucleotides, either in the single-stranded form or annealed to each other, and were incubated with increasing amounts of TraM (Fig. IV. 6). TraM was able to retard the double-stranded DNA in the mixture of the two oligonucleotides annealed together (Fig. IV. 6 **B**, lane 5) as well as the small amount of double-stranded DNA

Figure IV. 6

DNA retardation assay of the ColE1 *oriT* region synthetic oligonucleotides. The labeled fragments were incubated with TraM as described in Experimental Procedures and the mixtures were separated on 12% non-denaturing gels. Panel **A**. Single-stranded oligonucleotide $^{5'}$GGGTGTCGGGGCAGCCCATGACCC$^{3'}$. Panel **B**. Double-stranded oligonucleotide. Note that the sequence of the two oligonucleotides presented 5' to 3' are complementary. Panel **C**. Single-stranded oligonucleotide $^{5'}$GGGTCATGGCTGCCCCGACTCCC$^{3'}$. Lane 1, 10.0 pmole of synthetic DNA. The mobility of double-stranded DNA fragment and single-stranded DNA are indicated as well as the mobility of the DNA-protein complex (→). The asterisks indicate the mobility of the structures described in the Results section .


A
5' GGGTGTCGGGCGCAGCCATGACCC 3'
1 2 3 4 5
B
5' GGGTGTCGGGCGCAGCCATGACCC 3'
+
5' GGGTCATGGCTGCGCCCGACACCC 3'
1 2 3 4 5
C
5' GGGTCATGGCTGCGCCCGACACCC 3'
1 2 3 4 5
*
*
*
ssDNA
dsDNA
TraM (pmol)
-- .11 1.1 11 110
-- .11 1.1 11 110
-- .11 1.1 11 110

present in the samples of single-stranded oligonucleotides and indicated in Figure IV. 6 by asterisks. Because of the palindromic nature of the oligonucleotides, the DNA may have formed hairpin structures or concatamers, which would be substrates for TraM binding. Densitometric scanning of the autoradiogram showed that almost 100% of the complementary oligonucleotides annealed to form the double-stranded DNA. If we assume that TraM has four DNA binding sites per tetramer, then the molar ratio of DNA to protein is ≈1:40 in lane 5 of Figure IV. 6 **B**.

The double-stranded form of the oligonucleotides was also used in a competition assay of TraM binding to the *oriT* region of F. A DNA fragment from pLDLF13, containing all three TraM binding sites, was labeled at one end and incubated with an amount of TraM that would saturate the three binding sites. This mixture was incubated with increasing amounts of unlabeled double-stranded oligonucleotide and separated by electrophoresis on an 8% polyacrylamide gel. The molar ratio of *oriT* fragment to TraM protein was 1:2500 in every sample. Even when the double-stranded oligonucleotide was in a twenty fold excess compared to the amount of TraM, it was capable of only partially inhibiting the binding of TraM to the F *oriT* region (data not shown).

D. Discussion

F TraM is a small protein of 127 amino acids with a molecular weight of 14,507, as derived from the DNA sequence. The protein has an apparent molecular weight on SDS polyacrylamide gels of 10,950. This increased mobility is a common feature of other TraM proteins identified so far (Chapter II, Schwab *et al.*, 1991; Dempsey and Fee, 1990). Anti-TraM antiserum was prepared and used to localize the protein in cells carrying the whole F plasmid alone or together with a multicopy plasmid containing the *traM* gene. The protein was found mainly in the cytoplasmic fraction, in agreement with previous results for the TraM proteins of the F-like plasmids, pED208 and R1 (Chapter II, Schwab *et al.*, 1991). We were also able to detect large quantities of the protein in the inner membrane fractions of the strains carrying the *traM* gene on a multicopy plasmid. The presence of the F plasmid does not affect the partitioning of TraM although cells carrying chimeric plasmids containing the *oriT* region and TraM in the presence of F, have a greater amount of TraM in the inner membrane fraction than in the cytoplasm (data not shown). This suggests that a *tra* function recognizes the *oriT* region and affects the localization of TraM in the cells.

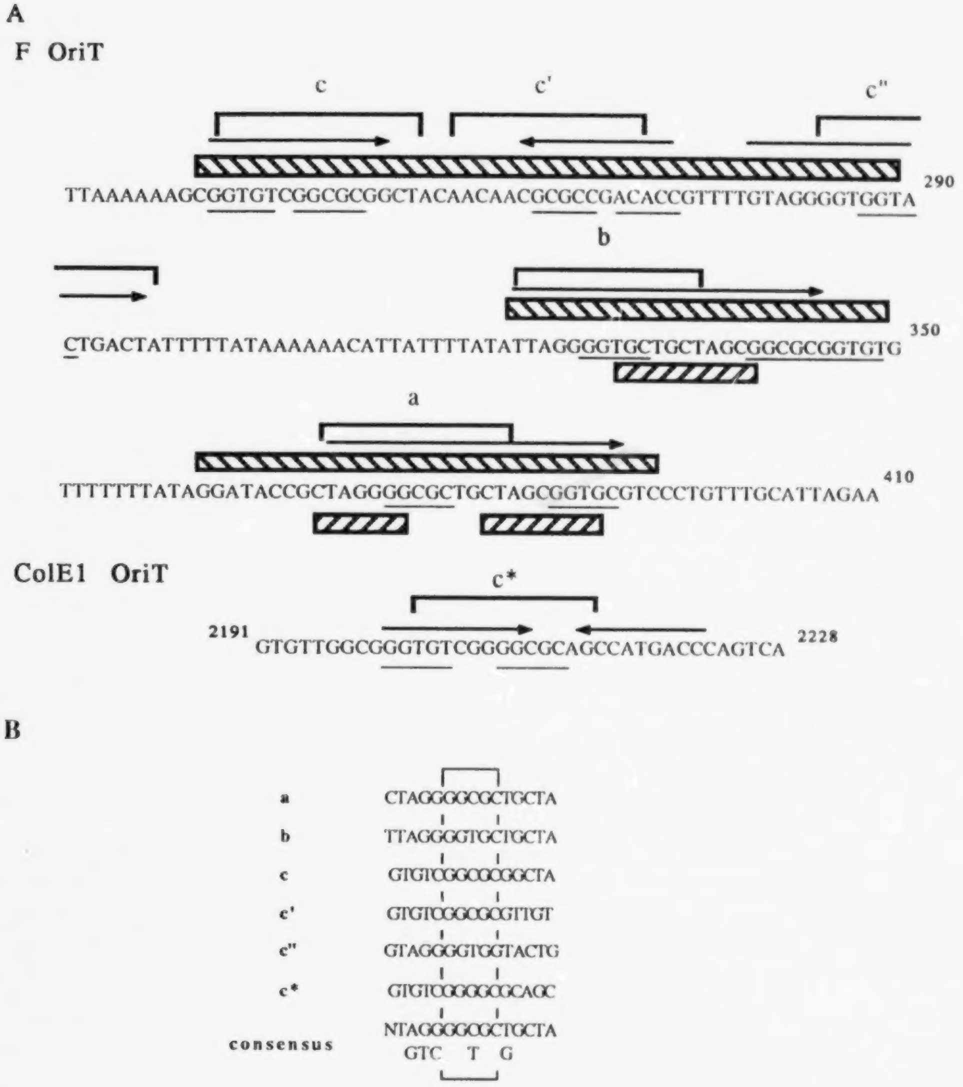
Electrophoretic mobility retardation assays were performed on the *oriT* regions of F and ColE1 incubated in the presence of F TraM. The *oriT* region of F showed three shifts while the ColE1 fragment showed one. An extra step was seen in both systems at high concentrations of TraM. We interprete this as aggregation of free TraM protein with TraM bound to the DNA, causing a mass action effect and electrophoretic mobility retardation of the DNA. It is interesting to note that similar amounts of TraM cause the formation of complex III in the F *oriT* fragment and complex I in the ColE1 *oriT* fragment, suggesting that TraM has a lower affinity for the ColE1 *oriT* region. This was further demonstrated by the inability of the double-stranded oligonucleotide from ColE1 to fully compete with the F *oriT* fragment for TraM.

The nature of the DNA-protein complexes was also analyzed using DNase I footprinting. Three major footprints were identified in the F OriT. The "footprints" encompass a 150 bp region including part of the *traM* promoter. The latter region was protected first, followed by the *oriT* region. Unlike the pED208 system (Chapter II) in which the interaction of the TraM protein with both strands of the DNA duplex was comparable, there is a remarkable difference in the affinity of the TraM protein for the two DNA strands in the F system. The different affinity for the two DNA strands may reflect a mechanism of interaction of F TraM with DNA which is different from pED208 TraM. TraM protects only the -35 region and to a lesser extent the -10 region of the *traM* promoter in the lower DNA strand. The localization of the protected regions within the regulatory region of the *traM* gene, as well as the order in which the binding occurred, suggests a role for TraM in the regulation of its own expression.

Sequence comparison of the regions protected by TraM identifies a pentameric motif of sequence GGC/TGC present 11 times within the F *oriT* region. Six have 1 or 2 non-conserved bases (Fig. IV. 7). Four of these motifs are spaced 12 bases apart within the protected regions suggesting that phasing along the DNA sequence is important. A fifteen base pair motif, repeated five times within the three footprints (Fig. IV. 7, panel **B**) and present in the ColE1 *oriT* region, was also identified. This motif is rich in G and is present on the upper strand four times out of five in the F OriT sequence. The pentamer GGPyGC is the core of this motif with the highly conserved G residues at positions 6 and 7 and a C residue at position 10.

Figure IV. 7

Summary of the TraM binding sites in the F *oriT* region. **A**. Sequences of the four protected regions (three from F and one from ColE1). The inverted repeats represent the secondary structures identified by Willetts and Wilkins (1984), whereas the direct repeats are from Thompson *et al.* (1984). The footprints are indicated by the striped boxes, on the top of the sequence for the upper strand and on the bottom for the lower strand; the open triangles indicate single base protection. The numbers beside the ColE1 fragment correspond to their position in the plasmid pBR322 beginning at the *Eco*RI site (Maniatis *et al.*, 1982). The TraM binding sites, corresponding to the 15 nucleotide motif discussed in the text, are identified by the horizontal brackets and named **a**, **b**, **c**, **c'**, **c"**. The core sequences of five bases described in the Discussion plus the mismatched sequences are underlined. **B**. The consensus sequence derived by a comparison of the DNA binding sequences for F TraM. The pentamer GGPyGC is boxed.

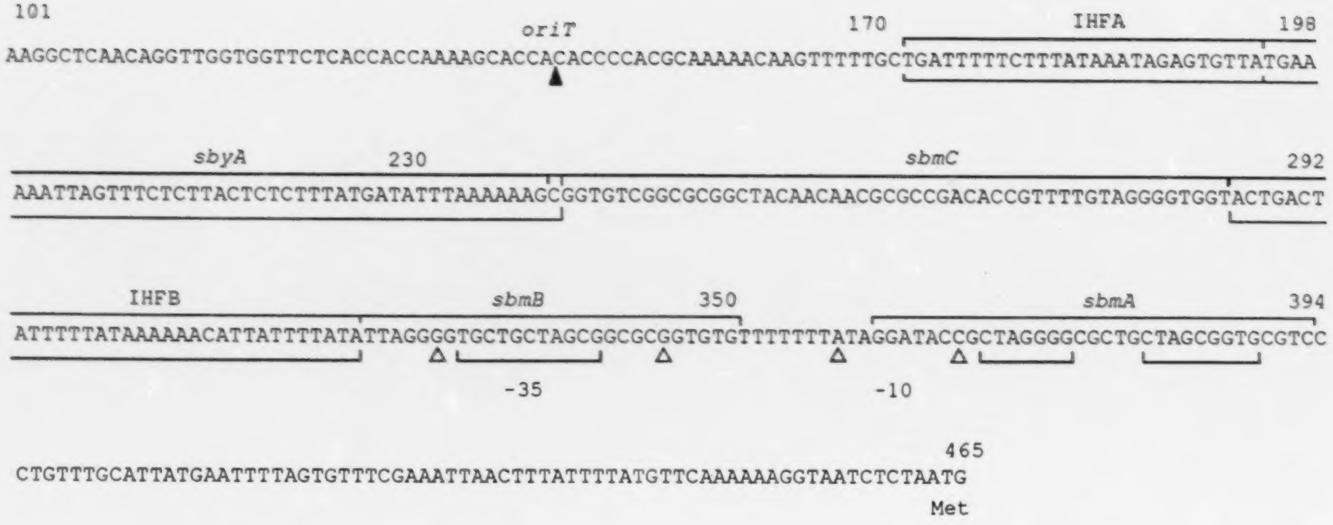

A
F OriT
c
c'
c"
TTAAAAAAGCGGTGTCGGCGCGGCTACAACAACGCGCCGACACCGTTTTGTAGGGGTGGTA
290
b
CTGACTATTTTTATAAAAAACATTATTTTATATTAGGGGTGCTGCTAGCGGCGCGGTGTG
350
a
TTTTTTTATAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCCCTGTTTGCATTAGAA
410
ColE1 OriT
c*
2191
GTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCA
2228
B
a
CTAGGGGCGCTGCTA
b
TTAGGGGTGCTGCTA
c
GTGTCGGCGCGGCTA
c'
GTGTCGGCGCGTTGT
c"
GTAGGGGTGGTACTG
c*
GTGTCGGGGCGCAGC
consensus
NTAGGGGCGCTGCTA
GTC T G

Factor or IHF (for review see Friedman, 1988; Yang and Nash, 1989) with the F *oriT* region *in vitro*. Two IHF binding sites were identified (Fig. IV. 8, IHF A and IHF B). The first site, which is closer to the nick site, was identified as a high affinity site, whereas the second site is weaker and is located between the two binding sites *sbmB* and *sbmC* (Tsai *et al.*, 1990) (Fig. IV. 8). Lahue and Matson (1990) characterized the gene product of *traY* and found that it binds to the *oriT* region (Fig. IV.8, *sbyA*). The TraY binding site is adjacent to the third TraM binding site (Fig. IV.8, *sbmC*), suggesting that the two proteins, TraY and TraM, might interact and in this way trigger the transfer process. TraI relaxes DNA by nicking at *oriT* with 50% efficiency in *in vitro* experiments but is unable to protect *oriT* DNA in DNase I footprinting experiments (Matson and Morton, 1991; Reygers *et al.*, 1991). Even though TraY is not necessary for nicking *in vitro* (Lahue and Matson, 1990), we suggest that TraI and TraY could interact in the nicking complex *in vivo* to selectively and efficiently nick the DNA at a site located almost 100 bases upstream from the TraY binding site, after a mating pair has been established (Fig. IV. 8, *oriT*).

We propose that TraM binds to the higher affinity sites, *sbmA* and *sbmB*, and represses its own expression. Upon formation of a stable mating pair, TraM could bind to *sbmC* and signal to the TraYI complex to nick the DNA and unwind it in preparation for transfer. This model is in agreement with the data presented by Fu *et al.* (1991); deletion analysis of the *oriT* region identified two sequences from nucleotide 237 to 285 and from nucleotide 325 to 360 that are essential for transfer, these sequences corresponded to *sbmC* and *sbmB*, respectively. The function of IHF would be to bend the DNA and bring into close proximity the nick site and the

Figure IV. 8

Summary of the interactions at the F *oriT* region. The sequence of the *oriT* region of F is shown with the positions where the *tra* gene products interact with DNA. The TraM binding sites (*sbmA*, *sbmB* and *sbmC*) are based on the work presented in this Chapter, the site of specific nicking (*oriT*) as well as the site of binding of TraY (*sbyA*) and IHF (IHF A and IHF B) proteins are based on published data (Reygers *et al.*, 1991; Tsai *et al.*, 1990; Lahue and Matson, 1990). The -35 region and -10 region are shown as published by Thompson and Taylor (1982).


101
oriT
170
IHFA
198
AAGGCTCAACAGGTTGGTGGTTCTCACCACCAAAAGCACCACACCCCACGCAAAAACAAGTTTTTGCTGATTTTTCTTTATAAATAGAGTGTTATGAA
sbyA
230
sbmC
292
AAATTAGTTTCTCTTACTCTCTTTATGATATTTAAAAAAGCGGTGTCGGCGCGGCTACAACAACGCGCCGACACCGTTTTGTAGGGGTGGTACTGACT
IHFB
sbmB
350
sbmA
394
ATTTTTATAAAAAACATTATTTTATATTAGGGGTGCTGCTAGCGGCGCGGTGTGTTTTTTTATAGGATACCGCTAGGGGCGCTGCTAGCGGTGCGTCC
-35
-10
465
CTGTTTGCATTATGAATTTTAGTGTTTCGAAATTAACTTTATTTTATGTTCAAAAAAGGTAATCTCTAATG
Met

site of interaction of TraYI with the DNA, and perhaps to affect superhelical density locally.

The characterization of the interaction of TraM with the F *oriT* region complements recent data on the characterization of other factors that play a role in the process of F transfer. Tsai *et al.* (1990) shows the interaction of Integration Host in the *oriT* region (Travers, 1990). Recent data published by Reygers *et al.* (1991) and Matson and Morton (1991) show that TraI has increased efficiency of nicking if the substrate DNA is negatively supercoiled. We are currently investigating the effects of supercoiling on the interactions of TraM with its binding sites, in order to elucidate the role played by TraM in initiating the transfer process and in the regulation of its own expression.

Willetts and Wilkins (1984) reported that the ColE1 *oriT* region has a sequence of dyad symmetry homologous to a sequence in the *oriT* region of F. They hypothesized the involvement of F *tra* genes in the mobilization of ColE1. Genetic studies performed on the ColE1 plasmid and the related plasmid CloDF13 (Willetts, 1980) showed that the *traM* function was not essential for plasmid mobilization. However, the nature of the mutation in *traM* has not been reported (it is not a nonsense mutant) and it could be that the (F*lac traM*102) mutants are still effective in mobilization. In this study, we have demonstrated the specific interaction of one of the *tra* genes with the *oriT* region of the mobilizable plasmid ColE1; however further characterization of the interaction of TraM and of other *tra* gene products with the ColE1 *oriT* as well as re-evaluation of the genetic studies, will be necessary to gain a better understanding of the mechanism of mobilization of non-conjugative plasmids.

E. Bibliography

Abo, T., Inamoto, S., Ohtsubo, E. (1991) Specific DNA binding of the TraM protein to the *oriT* region of plasmid R100. *J Bacteriol* **173:** 6347-6354.

Achtman, M., Willetts, N.S., and Clark, A.J. (1971) Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. *J Bacteriol* **106:** 529-538.

Bolivar, F., Rodriquez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H., and Falkow, S. (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. *Gene* **2:** 95-113.

Dempsey, W.B., Fee, B.E. (1990) Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol Microbiol* **4:** 1019-1028.

Dretzen, G., Bellard, M., Sassone-Corsi, P., Chambon, P. (1981) A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. *Anal Biochem* **112:** 295-298.

Finlay, B.B., Paranchych, W., Falkow, S. (1983) Characterization of conjugative plasmid EDP208. *J Bacteriol* **156:** 230-235.

Friedman, D.I. (1988) Integration host factor: a protein for all reasons. *Cell* **55:** 545-554.

Frost, L.S., Finlay, B.B., Opgenorth, A., Paranchych, W., and Lee, J.S. (1985) Characterization and sequence analysis of pilin from F-like plasmids. *J Bacteriol* **164:** 1238-1247.

Fu, F.Y.-H., Tsai, M.-M., Luo, Y., Deonier, R.C. (1991) Deletion analysis of the F plasmid *oriT* locus. *J Bacteriol* **173:** 1012-1020.

Ippen-Ihler, K. (1989) Bacterial Conjugation. In: *Gene Transfer in the Environment*, edited by Levy, S.B. and Miller, R.V. New York: McGraw-Hill Publishing Co., p. 33-72.

Jacobs, E., and Clad, A. (1986) Electro-elution of fixed and stained membrane proteins from preparative sodium dodecylsulfate polyacrylamide gels into a membrane trap. *Anal Biochem* **154:** 583-589.

Kingsman, A. and Willetts, N. (1978) The requirements for conjugal DNA synthesis in the donor strain during the F*lac* transfer. *J Mol Biol* **122:** 287-300.

Lahue, E.E., and Matson, S.W. (1990) Purified *E.coli* F-factor TraY protein binds *oriT*. *J Bacteriol* **172:** 1385-1391.

Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Matson, S.W., and Morton, B.S. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J Biol Chem* **266:** 16232-16237.

Schwab, M., Gruber, H., Hogenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol Microbiol* **5:** 439-446.

Studier, W.F., and Moffat, B. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. *J Mol Biol.* **189:** 113-130.

Reygers, U., Wessel, R., Müller H., and Hoffmann-Berling H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I direct against the transfer origin of the F factor. *EMBO J* **10:** 2689-2694.

Sambrook, J., Maniatis, T., Fritsch, E.F., and Irwin, N. (1989) *Molecular Cloning A Laboratory Manual* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Tabor, S., and Richardson, C.C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* **82:** 1074-1078.

Thompson, R., Taylor, L. (1982) Promoter Mapping and DNA Sequencing of the F Plasmid Transfer Genes *traM* and *traJ*. *Mol Gen Genet* **188**: 513-518.

Thompson, R., Taylor, L., Kelly, K., Everett, R., and Willetts, N. (1984) The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks. *EMBO J* **3**: 1175-1180.

Travers, A. A. (1990) Why bend DNA? *Cell* **60**: 177-180.

Traxler, B.A., and Minkley, E.G. Jr.(1988) Evidence that DNA helicaseI and *oriT* site specific nicking are both functions of the F TraI protein. *J Mol Biol* **204**: 205-209.

Tsai, M.-M., Fu, F.Y.-H., Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J Bacteriol* **172**: 4603-4609.

Voller, A., Bidwell, D., Huldt, G., and Engvall, E. (1974) A microplate method of ELISA and its application to malaria. *Bull W.H.O.* **51**: 209-221.

Willetts, N. (1980) Interactions between the F conjugal transfer system and CloDF13::Tna plasmids. *Mol Gen Genet* **180**: 213-217.

Willetts, N., and Skurray, R. (1987) Structure and Function of the F Factor and Mechanism of Conjugation. In: *Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology*, edited by Neidhardt, F.C., Ingraham, J.L., Magasanik, B., Low, K.B., Schaechter, M. and Umbarger, H.E. Washington: American Society For Microbiology, p. 1110-1133.

Willetts, N., Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol Rev* **48:** 24-41.

Yang, C.-C., Nash, H.A. (1989) The interaction of *E.coli* IHF protein with its specific binding sites. *Cell* **57:** 869-880.

CHAPTER V

Physico-chemical Characterization of the Signal Protein TraM from the Conjugative Plasmids F and pED208

A. Introduction

Among the IncF systems there are five classes of TraM. The prototypes of the five classes are: F, R100, R1, ColB4 (Willetts and Maule, 1985; Finlay *et al.*, 1986) and F_0*lac* (pED208 is the derepressed form of this plasmid;Chapter II). Members of the same class can complement each other whereas members of different classes cannot. TraM from four of the prototype conjugative plasmids was shown to bind to DNA sequences specific for each plasmid, located close to the *oriT* site (Chapter II, Chapter IV, Abo *et al.*, 1991, Schwab *et al.*, 1991). Among the five types of TraM, pED208 TraM is the least conserved, while F TraM is considered the paradigm. The purpose of this chapter is to characterize the physico-chemical properties of TraM from F and pED208 and compare the two proteins.

B. Experimental Procedures

1. Prediction of secondary structure using computer programs

The following computer programs were applied to the sequences of F and pED208 TraM proteins using the GeneWorks version 2.0 package (IntelliGenetics, Inc., Mountain View, CA) for Macintosh computers:

Garnier protein structure prediction. This algorithm predicts secondary structures for a protein sequence using the weight matrices developed by Garnier, Osguthorpe and Robson (1978). It predicts the likelihood that a particular residue will be found in an alpha helix, beta sheet, random coil or turn. For each amino acid in the sequence the program calculates four equations, one for each of the conformational states. The program then determines the most likely conformation by choosing the equation with the highest score. A diagram for each predicted conformational state as well as the composite diagram of the most likely conformation is produced.

Surface probability. This algorithm predicts the residues on the surface of a polypeptide by calculating the product of the surface accessibilities of each amino acid over a range of the sequence (Janin *et al.*, 1978).

2. Analytical ultracentrifugation

Analytical ultracentrifugation analysis was performed by Leslie Hicks in the laboratory of Dr. C. Kay (Department of Biochemistry). A Beckman Model E analytical ultracentrifuge equipped with electronic speed control and RITC temperature control was used for all runs. The concentration of TraM in the samples was determined using ultracentrifugal synthetic boundary runs with Raleigh interference optics (Babul and Stellwagen, 1969).

A measurement of the number of fringes crossed on the photographic plate using a Nikon Model 6 microcomparator, was related to the protein concentration using the average refractive index increment of 4.1 fringes mg^{-1} ml^{-1} determined for proteins by Babul and Stellwagen (1969).

Determinations of molecular weights were made using the conventional sedimentation equilibrium technique described by Chervenka (1970). Prior to running in the ultracentrifuge, the samples were dialysed against 0.1 M NaCl, 50 mM Tris-HCl pH 8.5, 0.1 mM disodium EDTA, 1mM dithiotreithol (DTT) (0.1 M NaCl/ TED buffer) for 48 hours. Fringe counts were then performed on the samples to determine the initial concentrations. One hundred µl of each sample was then loaded into a 12 mm double-sector, charcoal filled Epson cell equipped with sapphire windows. The sedimentation equilibrium runs were performed at 20°C for a minimum of 48 hours before equilibrium photographs were taken.

The weight average molecular weight of TraM was determined as a function of the initial concentration of the samples. The concentration of the protein (C) as a function of its distance (r) from the axis of rotation at equilibrium was measured using Raleigh interference optics. The concentration of protein at the meniscus, C_m, was calculated as follows:

$$C_m = C_o - \frac{r^2(C_b - C_m) - \int_{c_m}^{c_b} r^2 dc}{r^2_b - r^2_m} \qquad [1]$$

where C is the protein concentration at any distance, r, from the axis of rotation and the following subscripts indicate: o, value at time 0, m, value at the meniscus, b, value at the bottom of the cell. A direct measurement of protein concentration in terms of fringes, y, at any given value of r, was determined by adding the fringe displacement at point r to the concentration at the meniscus (expressed in fringes).

A plot of the natural logarithm of the protein concentration (in fringes, y) yielded a straight line, the slope of which, dlnc/dc^2 was used to determine the weight average molecular weight from the following equation at any position along the cell:

$$M_w(r)=\frac{2RT}{(1-\bar{v}\rho)\omega^2}\ \frac{d\ln C(r)}{dr^2} \qquad [2]$$

where R is the universal gas constant, T the experimental temperature in Kelvin, $\bar{v}$ is the protein partial specific volume, ρ is the density of the buffer at the temperature T, and ω is the angular velocity in radians per second. Point-average molecular weight calculations were carried out using a computer program written in APL language. The lnC versus r^2 data were fitted to a second degree polynomial equation using least squares techniques and the point-average molecular weights were calculated from the slope of the plot.

3. Circular dichroism

Circular dichroism analysis was performed by Kim Oikawa in the laboratory of Dr. C. Kay. Circular dichroism (CD) measurements were done on a Jasco J-720 spectropolarimeter (Jasco Inc., Easton, MD) interfaced to an Epson Equity 386/25 and controlled by Jasco software. The thermostated cell holder was maintained at 25°C with a Lauda RMS circulating water bath (Lauda, Westbury, NY). The instrument was routinely calibrated with ammonium d-(+)-10 camphor sulfonate at 290.5 nm and 192 nm, and with d-(-)-pantoyllactone at 219 nm. Each sample was scanned ten times and noise reduction applied to remove the high frequency noise before calculating molar ellipticities. The voltage of the photomultiplier was kept below 500 V to prevent distortion of the CD spectrum. The dimensions of the cells used for the measurements at wavelengths below 250 nm were 0.01 and 0.05 cm, whereas a 1 cm microcell was used for the aromatic region (250-320 nm).

The circular dichroic spectrum of a protein arises from a differential absorption of left and right circularly polarized light at any given wavelength, it can be defined as $\Delta\varepsilon=\varepsilon_L-\varepsilon_R$, where $\Delta\varepsilon$ is the circular dichroism and ε_L and ε_R are absorption coefficents of left and right circularly polarized light, respectively. As the emergent light is elliptically polarized, it is conventional to express it in terms of mean residue ellipticity $[\theta]_\lambda$ defined by the equation:

$$[\theta]_\lambda = \frac{MRW \cdot \theta_{obs}}{100 \cdot l \cdot c} \qquad [3]$$

where MRW is the mean residue weight corresponding to 116.024 for TraM (calculated from the sequences in Chapter II and Thompson and Taylor, 1982), θ_{obs} is the observed ellipticity value at the wavelength of interest, l is the pathlength of the cell in dm and c is the concentration of the protein in g cm^{-3}. The unit for molar ellipticity was degree centimeter squared per decimole. Protein concentration ranged from 0.5 to 2.38 mg ml^{-1}.

The CD spectra of proteins in the far UV region (190-230 nm) depend on their secondary structures. The ellipticity of a sample at any given wavelength can be expressed as the sum of the contribution from α-helix, β-sheet and random coil:

$$[\theta]_\lambda = f_\alpha\,[\theta]_{\alpha,\lambda} + f_\beta\,[\theta]_{\beta,\lambda} + f_{c,\lambda}\,[\theta]_{c,\lambda} \qquad [4]$$

where f_α, f_β, and f_c are the predicted fractions of α-helix, β-sheet and random coil respectively and $[\theta]_{\alpha,\lambda}$, $[\theta]_{\beta,\lambda}$ and $[\theta]_{c,\lambda}$ are the values that represent the known molar ellipticity of each structural configuration at any given wavelength. The relative intensity is based on the CD spectra of 5 proteins that have had their structures previously determined by X-ray crystallography (Chen *et al.*, 1974).

The fraction of α-helix and β-sheet was also determined using a program developed by Provencher and Glockner (1981) which analyzes CD spectra as a sum of spectra of 16 proteins whose structures have been determined by X-ray crystallography. Furthermore, the method includes a fourth structural configuration, the β-turn. The Contin I program (from the European Molecular Biology Laboratory) for applying this method was also used to analyze changes in the secondary structural content of the proteins after addition of metal ions. The input to the program was the molar ellipticities in 1 nm intervals from 190 to 240 nm.

The near UV CD spectrum (250-320 nm) of most proteins derives mainly from the aromatic residues tryptophan, tyrosine and phenylalanine. They possess a low level of CD when existing as free amino acids in solution. However, the Cotton effects can be greatly enhanced if these chromophores are found nearby asymmetric structures such as α-helix and β-forms and also upon interaction with other aromatic residues (Strickland, 1974).

C. Results

1. Sequence comparison of the TraM proteins from the conjugative plasmids F and pED208

The percentage of direct and conserved homology among the five prototypes of TraM, compared to F TraM, is reported in Table V. 1. It shows the high degree of homology among the five prototypes with the exception of pED208 TraM, which shows 41.8 % identity and 57.5 % conserved homology to F TraM. The amino acid sequences of the two proteins were compared as shown in Figure V. 1, where the identical (conserved) residues are indicated with a vertical bar. The amino terminal end from residue 1 to residue 35 is characterized by the presence of a cluster of basic residues, whereas the carboxy terminal end from residue 100 to 127 has a high content of acidic residues, with pED208 TraM having six acidic residues of the last seven. The central region (residue 32 to 99) is characterized by non-polar residues. The only cysteinyl residue in the polypeptides is located in the central region of both proteins, at positions 73 in F TraM and 79 in pED208 TraM.

In order to identify homology with other known proteins, the amino acid sequences of TraM from F and pED208 were compared to the PC/GENE data bank (PC/GENE: the nucleic acid and protein sequence analysis software system. A. Bairovoch, University of Geneva (TM) IntelliGenetics Inc, Mountain View, CA). No region of homology was detected with the sequences contained in the data bank.

Table V. 1

Amino Acid Sequence Homology Among the Five Prototypes of TraM

Plasmid	Direct homology (%)	Conserved homology (%)
F	100	100
ColB4	81.4	88.2
R100	88.9	93.7
R1	89.7	95.2
pED208	41.8	57.5

```
F         MAKVNLYISNDAYEKINAIIEKRRQEGAREKDVSFSATASMLLELGLRVHEAQMERKESAFNQTEFN
          | ||| || | ||| |  ||  |||||| | ||| |   ||||||||||  || ||||  ||| |||
pED208    MPKIQTYVNNNVYEQILDLVTIRKQEGI-E-EASLSNVSSMLLELGLRVYMIQQEKREGGFNQMEYN

          KLLECVVKTQSSVAKILGIESLSPH--VSGNSKFEYANMVEDIREKVSSEMER-FFPKNDDE
          || || | |       ||   |     ||||  ||||     | ||   |    |||  |||
          KLMLENVSRVRAMCTEILKMSVLNQESIASGN--FDYAVIKPAI-DKFAREQVSIFFPDDEDDQE
```

Figure V. 1

Amino acid sequence alignment of the TraM proteins from the conjugative plasmids F and pED208. One letter code was used for the amino acid residues and the conserved residues are connected by a vertical line. The sequences are from Thompson and Taylor (1982) and Chapter II.

2. Prediction of secondary structure

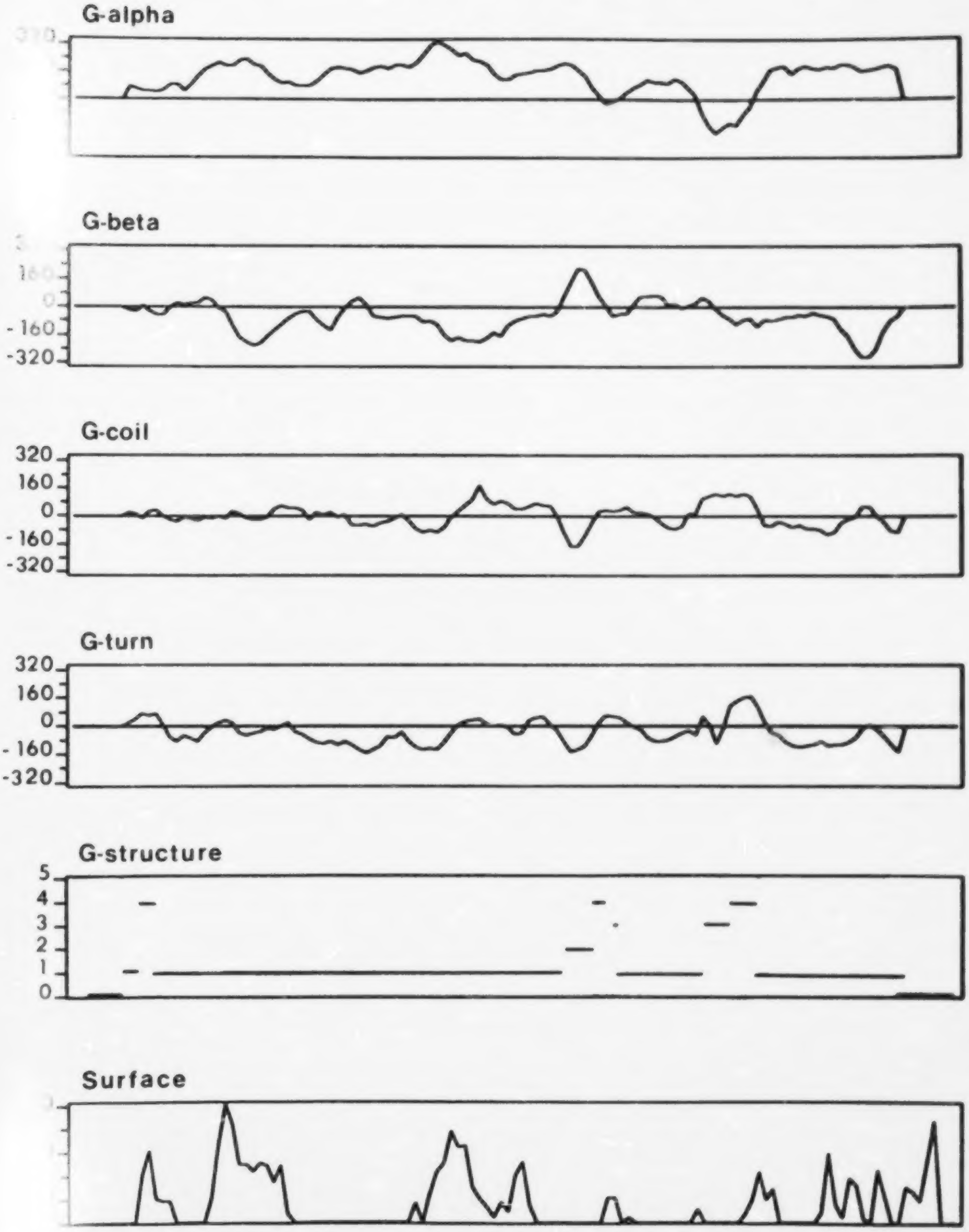
The amino acid sequences of the TraM proteins were analyzed for prediction of secondary structure using several modified versions of the Chou-Fasman method (Chou and Fasman, 1978). Since the proteins show DNA-binding activity (Chapter II and IV), the sequences were analyzed for potential helix/turn/helix structure, which is a well characterized folding motif of many prokaryotic DNA-binding proteins.

The sequences were analyzed by GeneWorks 2.0 computer programs. The comparative studies of the characteristics of the TraM proteins; secondary structure prediction using the Garnier (Garnier *et al.*, 1978) and Chou-Fasman methods, surface probability (Janin *et al.*, 1978) and chain flexibility (Karplus and Schulz, 1985) profiles, are shown in Figures V. 2 and V. 3. The predicted α-helical content of both proteins was high and a few regions of β-turn were also predicted as shown by the composite diagram (Fig. V. 2 and V. 3, G structure); the F TraM protein showed three helices, the first one between residues 13 and 63, the second between residues 80 and 92 and the third between residues 98 and 120 (Fig. V. 2), whereas the pED208 TraM protein showed five α-helices (Fig. V. 3).

The surface probability and chain flexibility plots were obtained for the proteins. The profiles showed both proteins to be flexible (data not shown) with three main regions exposed on the surface of the proteins corresponding to a basic region at the amino terminal end (9-32), the charged region in the middle of the polypeptide (44-65) and the acidic region at the carboxy terminal end (101-127) (Fig. V. 2 and V. 3, bottom graph).

Figure V. 2

Prediction of the secondary structure of the F TraM protein using the Garnier method (Garnier *et al.*, 1978) and the surface probability plot (Janin *et al.*, 1978). The four parameters of the Garnier method are indicated: G-alpha, α-helical structure, G-beta, β-sheet structure, G-coil and G-turn, coil and turn conformations. G-structure, composite diagram of the predicted conformations, where level 1 represents the α-helix, level 2 the β-sheet, level 3 the coil and level 4 the turn conformations. The surface plot shows the probability of a stretch of amino acids being on the surface .



G-alpha

G-beta

G-coil

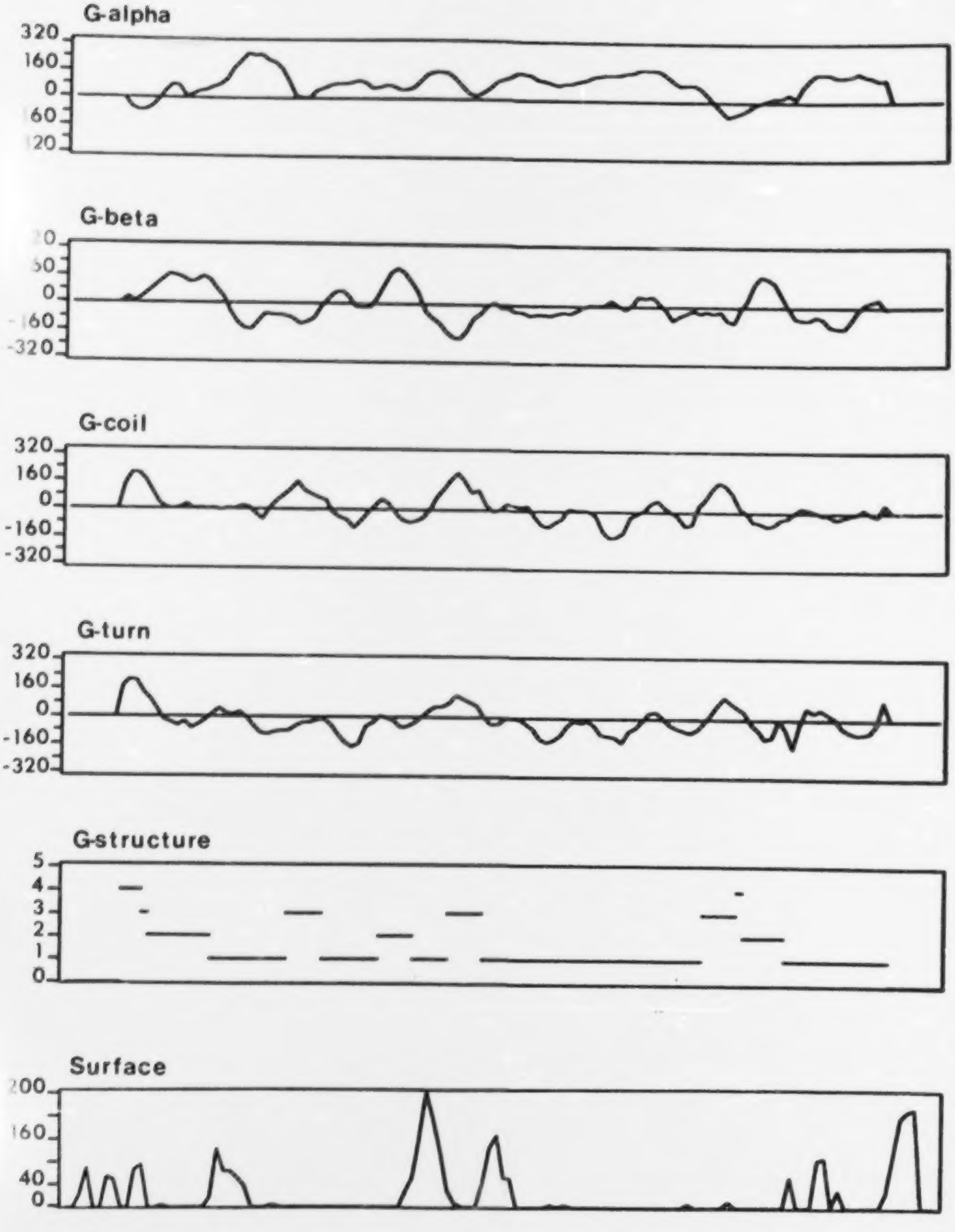
G-turn

G-structure

Surface

Figure V. 3

Prediction of the secondary structure of the pED208 TraM protein using the Garnier method (Garnier *et al.*, 1978) and the surface probability plot (Janin *et al.*, 1978). The four parameters of the Garnier method are indicated: G-alpha, α-helical structure, G-beta, β-sheet structure, G-coil and G-turn, coil and turn conformations. G-structure, composite diagram of the predicted conformations, where level 1 represents the α-helix, level 2 the β- sheet, level 3 the coil and level 4 the turn conformations. The surface plot shows the probability of a stretch of amino acids being on the surface.


G-alpha
320
160
0
160
120
G-beta
0
-160
-320
G-coil
320
160
0
-160
-320
G-turn
320
160
0
-160
-320
G-structure
5
4
3
2
1
0
Surface
200
160
40
0

3. Hydrodynamic characterization of TraM

The final chromatographic purification step of the TraM samples using the Blue Sepharose column (Chapter IV) yielded higher amounts of purified protein compared to the DNA-cellulose affinity chromatography (Chapter II), thus the proteins were purified for the physical-chemical characterization according to the protocol in Chapter IV. Since both proteins showed anomalous reactivity to the Lowry assay the final concentration of the purified samples were measured using absorbance at 280 nm or fringe counts.

Sucrose gradient centrifugation was used to determine the molecular weight of the pure proteins. Fifty µg aliquots of the purified TraM proteins were loaded on 5% to 20% sucrose gradients in TED buffer pH 6.4, and centrifuged at 140,000g for 20 hr at 4°C. As internal molecular weight standards, 10 µl of 1 mg ml^{-1} each of lysozyme, DNase I, Vit B12 and blue dextran were centrifuged together with the TraM samples. After ultracentrifugation, 200 µl fractions were collected, the refraction index of each fraction was measured using a refractometer (Bausch and Lomb). The samples were dialysed overnight against double-distilled water and dried before loading onto 15% denaturing polyacrylamide gels (SDS-PAG). Electrophoresis was performed at 30 mA at room temperature and the protein bands were visualized by silver staining. There was no detectable difference between the sedimentation profiles of the TraM proteins, and the TraM band appeared at sucrose concentrations between 13% and 16%; the internal standard DNase I was detected at similar sucrose concentrations, suggesting a molecular weight of about 60,000 for the TraM proteins (Fig. V. 4). Since Vit B12 and blue dextran are not proteins, they were not visualized by the silver staining. The calculated molecular weights of the

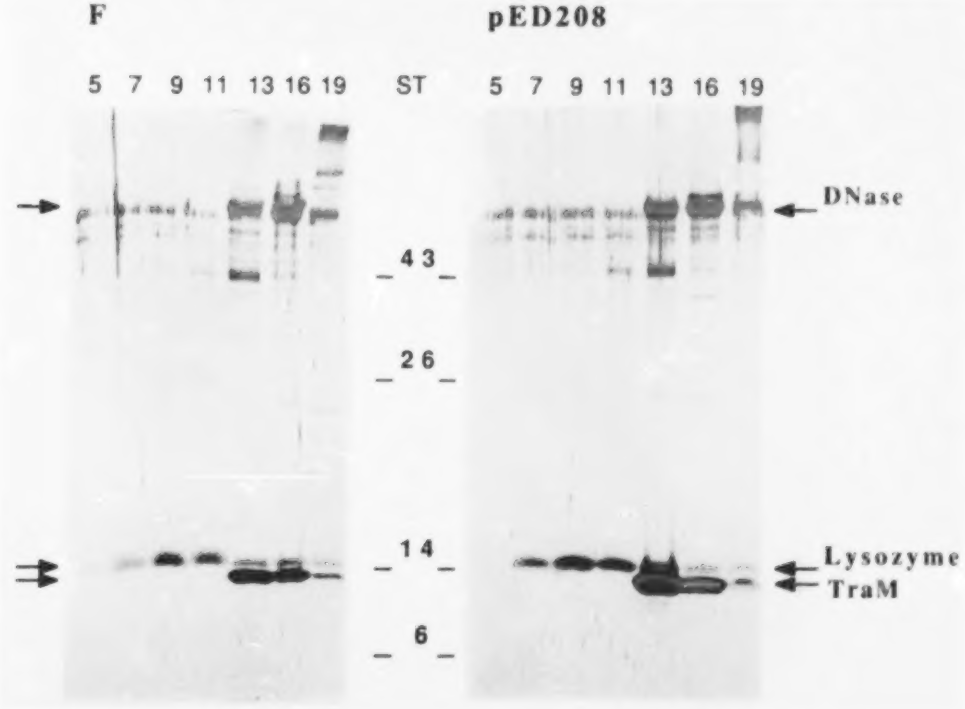

F
pED208
5 7 9 11 13 16 19
ST
5 7 9 11 13 16 19
DNase
43
26
14
6
Lysozyme
TraM

Figure V. 4

Silver staining of the sucrose gradients fractions separated by 15% SDS polyacrylamide gel electrophoresis. The internal molecular weight standards are indicated as well as the bands corresponding to the TraM proteins.The sucrose concentration (%) of each sample is indicated on the top of the gel. St, molecular weight standards are indicated in kilodaltons.

TraM proteins from their amino acid compositions were 14,507 and 14,542 for F TraM and pED208 TraM, respectively. Thus both proteins appeared to be homotetramers.

To investigate the effect of pH on the quaternary structure of the proteins, sucrose gradients were performed at pH 7.9 and pH 8.8. The results were identical to the gradients performed at pH 6.4 and suggested that the protein is a stable tetramer within the range of pH values used (data not shown).

Low speed sedimentation equilibrium studies were carried out for both proteins. For pED208 TraM a molecular weight of 61,870 was calculated at an initial loading concentration of 2.18 mg ml^{-1} in 0.1 M NaCl/TED buffer (Fig. V. 5). This value corresponded to the one obtained from sucrose gradients and supported the tetrameric structure configuration. The F TraM sample showed aggregation and molecular weights ranging from 40,000 to 120,000 were obtained. To obtain a more accurate molecular weight determination of F TraM, the sample was treated with 4M guanidine hydrochloride. After the removal of the denaturant the sample was dialysed to equilibrium against fresh buffer and analytical ultracentrifugation was performed. The plot of the natural logarithm of the protein concentration measured in fringes (y) versus r^2 showed a non linear profile and the molecular weights were calculated to range between 30,000 and 90,000; this suggested the presence of a mixture of molecular forms ranging from the dimer to the hexamer (Fig. V. 6).

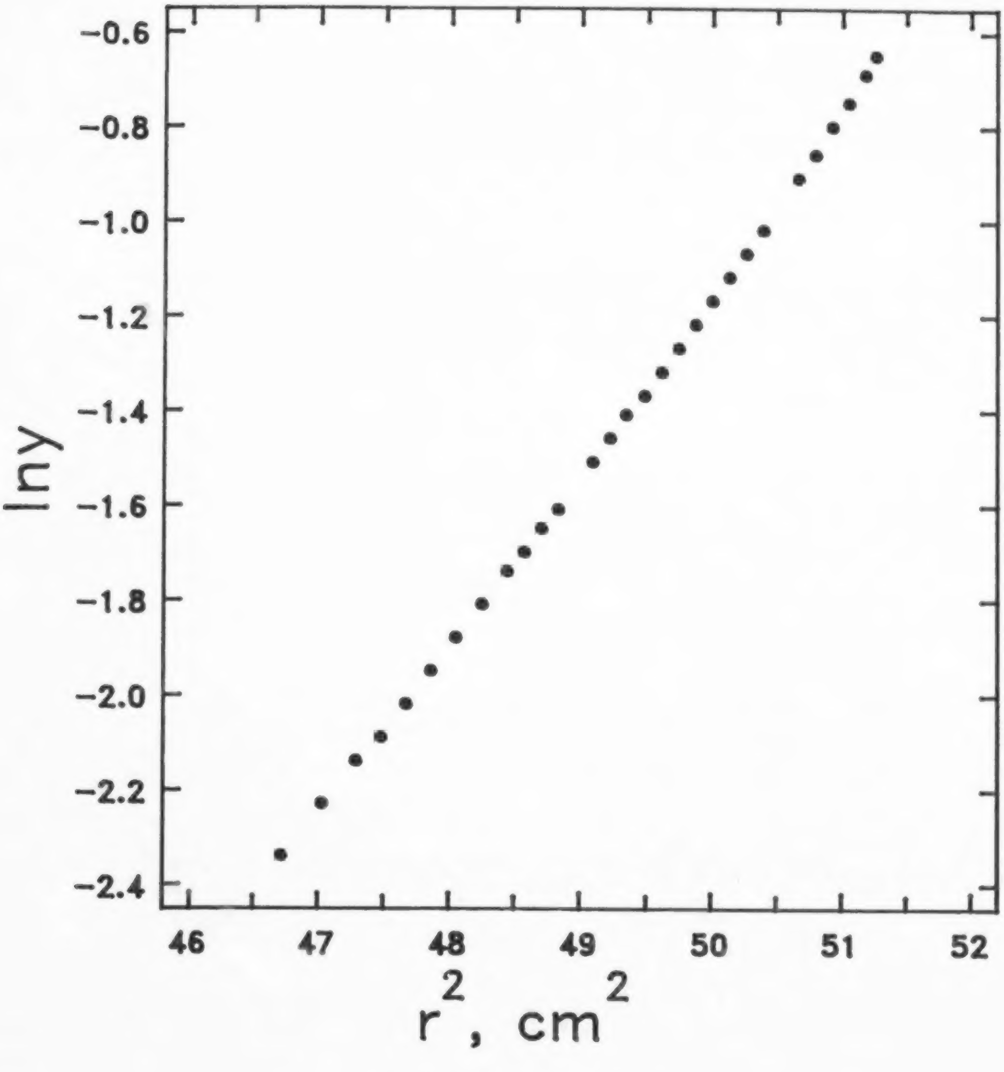


Figure V. 5

Determination of the weight average molecular weight of the pED208 TraM. The weight average molecular weight was determined from the slope of the plot using equation [2]. lny(r), the natural logarithm of the refraction index of the sample at the distance r from the axis of rotation (measured in cm).

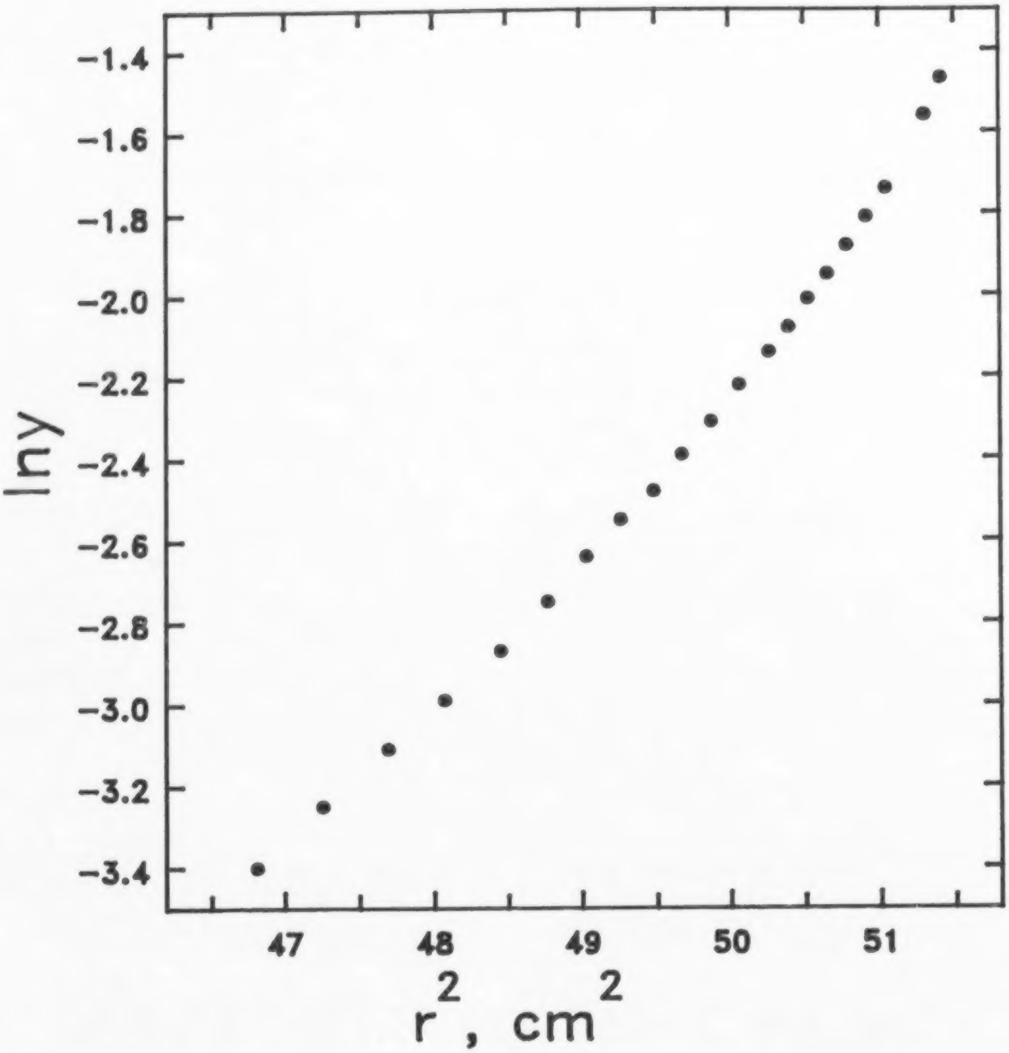


Figure V. 6

Determination of the weight average molecular weight of the F TraM. The weight average molecular weight was determined from the slope of the plot using equation [2]. lny(r), the natural logarithm of the refraction index of the sample at the distance r from the axis of rotation (measured in cm).

4. Spectrophotometric characterization of the TraM proteins

The ultra-violet absorption spectrum of TraM proteins (Fig. V. 7) showed a peak at 276 nm with a shoulder at 283 nm, both of which are attributed to the tyrosine chromophores. The fine structures in the spectrum near 260 nm induced by the phenylalanine residues were less prominent in pED208 TraM than F TraM since the phenylalanine to tyrosine ratio is one in pED208 TraM and two in F TraM.

The molar extinction coefficents of the proteins were determined to be 0.31 and 0.49 for F TraM and pED208 TraM, respectively, which is in agreement with the values of 0.29 and 0.51 calculated on the basis of the number of tyrosine residues per monomer.

The secondary structure of TraM was examined using far ultra-violet circular dichroism (UV-CD). The TraM spectra were typical of proteins containing significant amounts of α-helix. The CD spectra of both types of TraM were very close to that of a polypeptide chain in the α-helical conformation. The α-helix dichroic band normally found at 222 nm was shifted to 220 nm (Fig. V. 8). The measured ellipticity values at this wavelength were -24,410 and -21,030 degrees cm^2 $dmol^{-1}$ for F TraM and pED208 TraM, respectively.

Both proteins possess appreciable amounts of α-helix, as revealed by the CD measurements, although they differ in their absolute quantities of types of secondary structures. Reverse plot of the CD spectra measurements showed that F TraM was 79% α-helical and 21% random coil, whereas pED208 TraM was 66% α-helical, 16% β-sheet and 19% β-turn.

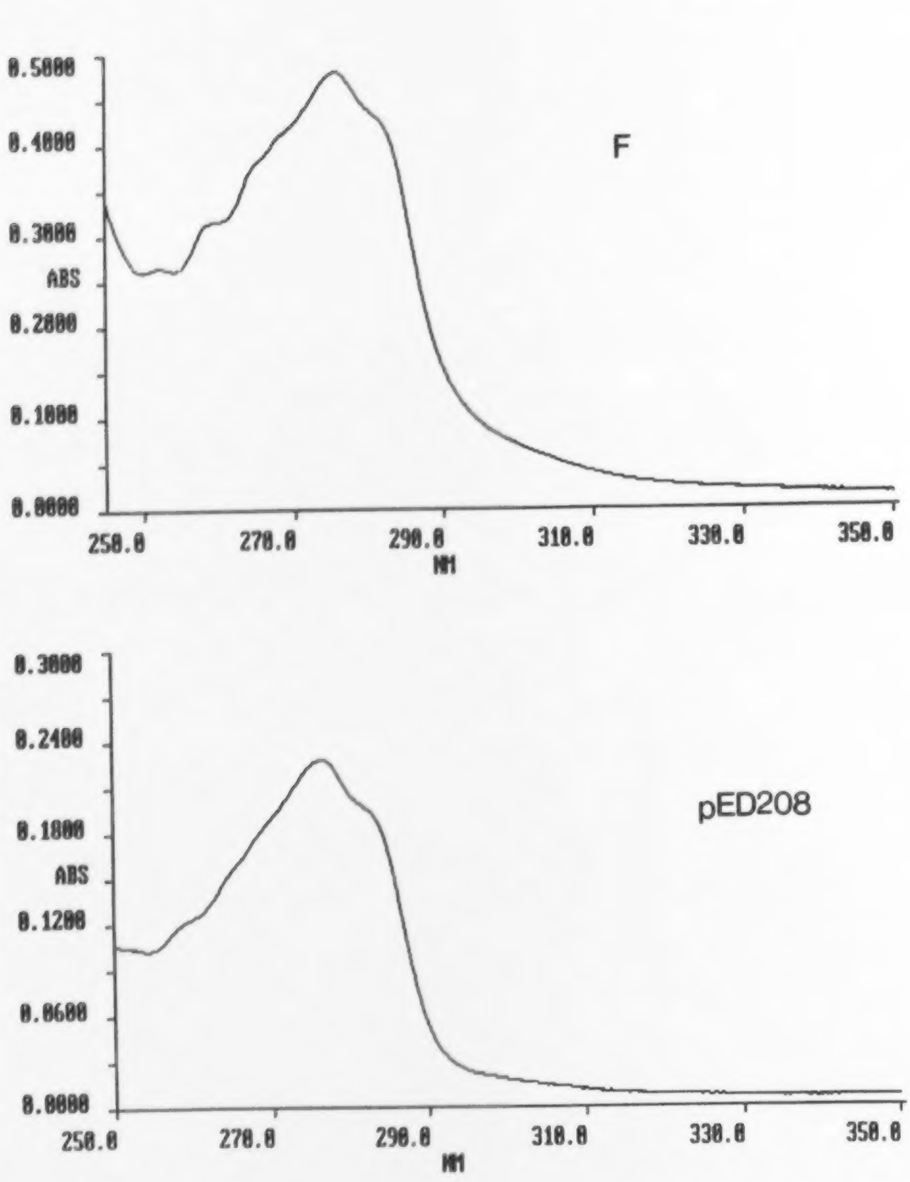
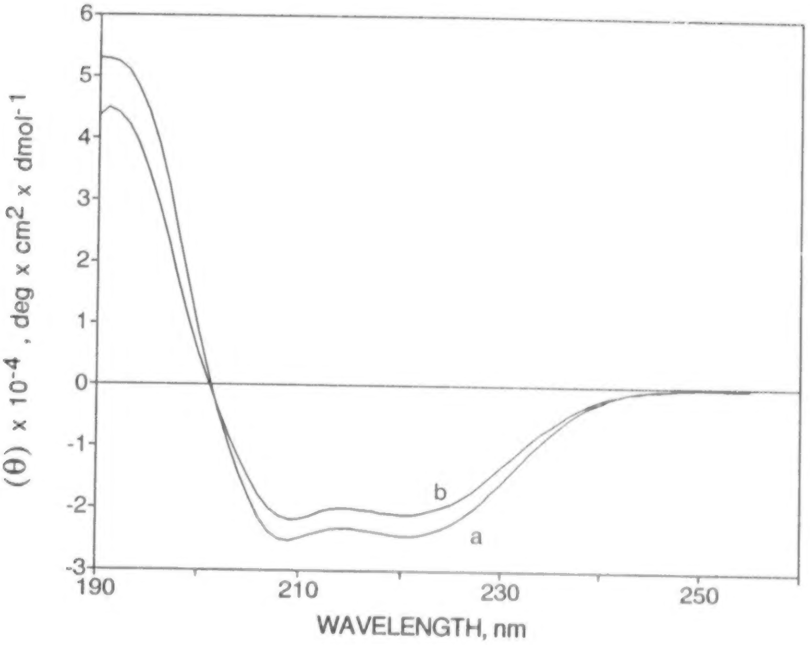


Figure V. 7

Ultra-violet scanning of the purified TraM proteins. The concentration of protein was 0.7 and 1.0 mg ml^{-1} for F and pED208 TraM, respectively.



. Figure V. 8

Far ultra-violet circular dichroism spectra of TraM proteins. The spectra were obtained at 0.71 mg ml^{-1} (a) of F TraM and 0.52 mg ml^{-1} (b) of pED208 TraM. The values of ellipticity are in degree cm^{2} dmol^{-1}.

The near UV CD spectra for the aromatic residues showed a similar profile for the two proteins although the absolute values of ellipticity differed (Fig. V. 9).

5. Immunodetection

Polyclonal antibodies against the F TraM protein were used in immunoblot assays of the purified TraM proteins of F and pED208 to detect the presence of common epitopes between the two proteins.

Seven μg of each purified protein was electrophoresed on a 15% polyacrylamide SDS gel and transferred to nitrocellulose membrane according to the protocol in Chapter II. A 1:100 dilution of α-F TraM antisera was incubated with the nitrocellulose bound proteins and the binding was detected with ^{125}I-labeled protein A. pED208 TraM cross-reacted weakly with the antibodies, as shown in Figure V. 10, suggesting the presence of common epitopes between the two proteins. The extra bands in the F TraM lane corresponded to the molecular weight of the dimer of F TraM.

D. Discussion

One of the properties of the TraM protein *in vitro* is to bind the *oriT* region as shown in the F-like systems R100 (Abo *et al.*, 1991), R1 (Schwab *et al.*, 1991), pED208 (Chapter II) and F (Chapter IV). Each TraM protein binds to a different sequence in the homologous *oriT* region. Neither one of the analyzed TraM proteins shows the structural motif α-helix-turn-α-helix, typical of DNA binding proteins,

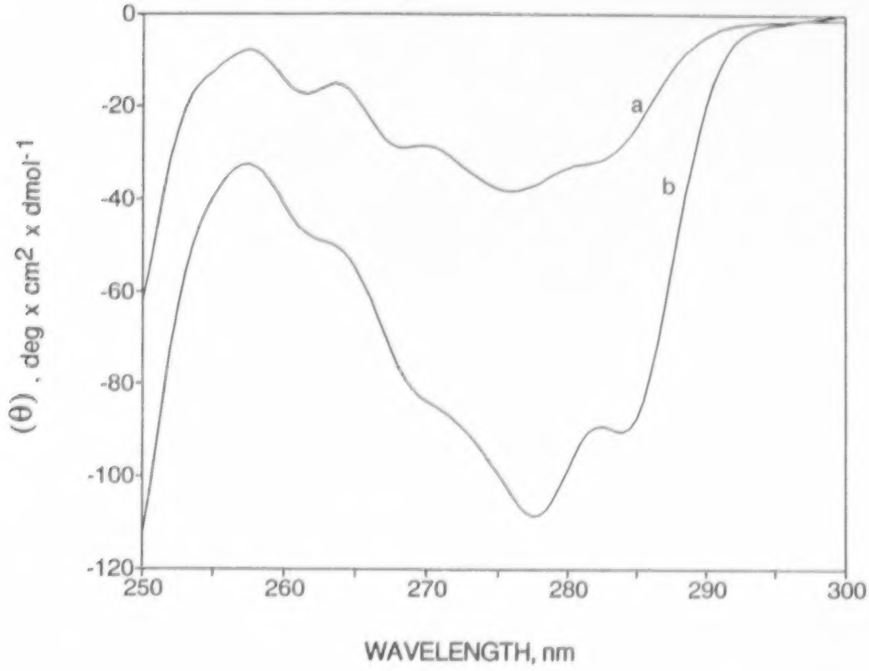


Figure V. 9

Near ultra-violet circular dichroism spectra of F and pED208 TraM. The concentration of F TraM was 2.38 mg ml^{-1} (a) and of pED208 TraM was 1.3 mg ml^{-1} (b). The values of ellipticity are in degree cm^2 $dmol^{-1}$.

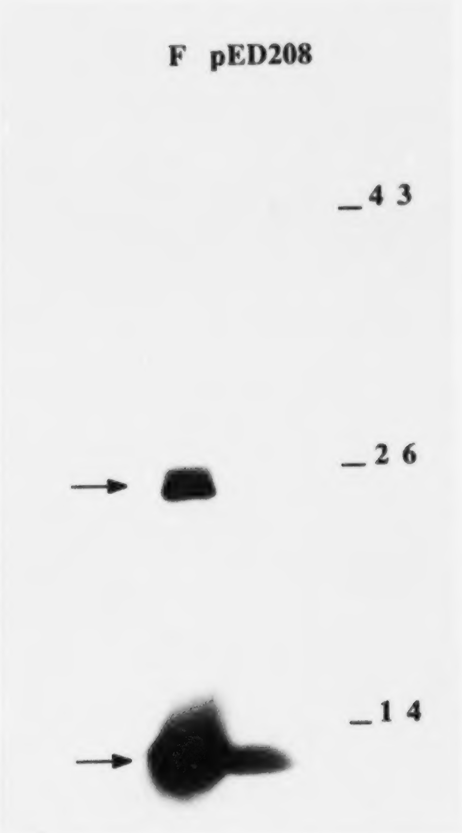


Figure V. 10

Immunoblot of the purified F and pED208 TraM proteins incubated with anti-F TraM antisera. The molecular weight standards, shown on the right-hand side of the figure, are expressed in kilodaltons. The mobility of the TraM proteins as well as of the dimer of F TraM are indicated.

although a few regions of α-helical structures interrupted by a turn or a coil conformation were predicted for both pED208 and F TraM proteins.

Studies done using minicells have shown TraM in the membrane of the cell (Achtman *et al.*, 1979; Kennedy *et al.*, 1977), whereas over-expressed TraM cloned into multicopy plasmids is recovered from the soluble cytoplasmic fraction (Chapter II and IV, Schwab *et al.*, 1991). Immunological characterization of TraM in cells carrying the whole conjugative plasmid found the protein in the cytoplasm with a small amount associated with the inner membrane (Chapter II and IV). The hydrophilicity plots, RAOARGOS (Argos *et al.*, 1982), SOAP (Kyte and Doolittle, 1982), HELIXMEM (Eisenberg, 1984) programs for the prediction of smoothed helix profile show profiles of very soluble proteins with no trans-membrane regions. This suggests that TraM may interact with other *tra* functions at the membrane level and together act as "sensor of mating establishment" complex. Upon stabilization of the mating pair, TraM may change the affinity of the complex for the DNA thus initiating the transfer process. One candidate for the sensor function is the nicking complex TraYI (Everett and Willetts, 1980). The TraI protein from F, which is DNA helicase I (Abdel-Mohem *et al.*, 1983), has been shown to nick *oriT* (Matson and Morton, 1991; Reygers *et al.*, 1991) but with low efficiency, and it does not bind to the *oriT* region. When present *in vitro*, TraY fails to increase the efficiency of the nicking reaction (Matson and Morton, 1991). However, the TraY binding site on the DNA (Lahue and Matson, 1990) is adjacent to one of the TraM binding sites (Chapter IV) suggesting that TraM may interact with TraY, thus stabilizing the TraYI complex, promoting efficient nicking and triggering the transfer process.

Both sucrose gradients and sedimentation equilibrium studies indicate TraM as a tetramer of the same subunit. The discrepancy between the values of molecular weights obtained for F TraM in the two techniques is attributed to the high concentration of protein (7.3 mg ml^{-1}) in the stock solution from which the sample used in the analytical centrifugation was derived by dilution.

The UV CD spectroscopy is a sensitive probe of the secondary structure of the TraM proteins. Both proteins show high α-helical content, suggesting a compact structure. The tyrosine residues appear to be in an asymmetric environment as detected by the near UV CD. Ca^{2+} and Mg^{2+} did not cause changes in the CD spectra of the proteins (data not shown) and conformational changes upon binding to DNA could not be detected because they were covered by the peak of maximum absorbance of the DNA at 260 nm (Chapter III).

In conclusion, the data presented in this chapter indicate that pED208 and F TraM share extensive sequence homology, similar predicted secondary structures and the same tetrameric subunit arrangement, although they bind to different DNA sequences and cannot complement each other. This suggests that the overall shape of the TraM protein as well as the mechanism of interaction with the *oriT* region, has been conserved during evolution among the IncF plasmids and a range of variation in the amino acid sequences of the two proteins can account for the observed diversity.

E. Bibliography

Abdel-Monem, M., Taucher-Scholz, G. and Klinkert, M. Q. (1983) Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. *Proc. Natl. Acad. Sci. U.S.A.* **80**: 4659-4663.

Abo, T., Inamoto, S. and Ohtsubo, E. (1991) Specific DNA Binding of the TraM Protein to the *oriT* Region of Plasmid R100. *J. Bacteriol.* **173**: 6347-6354.

Achtman, M., Manning, P. A., Edelbluth, C. and Herrlich, P. (1979) Export without proteolytic processing of inner and outer membrane proteins encoded by the F sex factor *tra* cistrons in *Escherichia coli* minicells. *Proc. Natl. Acad. Sci. U.S.A.* **76**: 4837-4841.

Argos, P., Rao, J. K. M. and Hargrave, P. A. (1982) Structural prediction of membrane-bound proteins. *Eur. J. Biochem.* **128**: 565-575.

Babul, J. and Stellwagen, E. (1969) Measurement of protein concentration with interference optics. *Anal. Biochem.* **28**: 216-221

Chen, Y.H., Yang, J.T., Chau, H.K. (1974) Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry.* **13**: 3350-3359

Chervenka, C.H. (1970) A manual of methods for the analytical ultracentrifuge, Spinco Division of Beckman Instruments, Palo Alto.

Chou, P.Y. and Fasman, G. D. (1978) Empirical predictions of protein conformations. *Annu. Rev. Biochem.* **47**: 251-276.

Eisenberg, D. (1984) Three-dimensional structure of membrane and surface proteins. *Annu. Rev. Biochem.* **53**: 595-623.

Everett, R. and Willetts, N. (1980) Characterisation of an in vivo system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136**: 129-150.

Finlay, B.B., Frost, L. S. and Paranchych, W. (1986) Origin of transfer of IncF plasmids and nucleotide sequences of the type II *oriT*, *traM*, and *traY* alleles from ColB4-K98 and the type IV *traY* allele from R100-1. *J. Bacteriol.* **168**: 132-139.

Garnier, J., Osguthorpe, D. J. and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structures of globular proteins. *J. Mol. Biol* **120**: 97-120.

Ippen-Ihler, K.A. and Minkley, E. G. Jr. (1986) The conjugation system of F, the fertility factor of *Escherichia coli*. *Annu. Rev. Genet.* **20**: 593-624.

Janin, J., Wodak, S., Levitt , M. and Maigret, B. (1978) Conformation of amino acid side-chains in proteins. *J. Mol. Biol.* **125**: 357-386.

Karplus, P.A. and Schulz, G. E. (1985) Prediction of chain flexibility in proteins. Naturwissenschlaften **72**: 212-213.

Kennedy, N., Beutin, L., Achtman, M.,Skurray, R., Rahmsdorf, U. and Herrlich, P. (1977) Conjugation proteins encoded by the F sex factor. *Nature* **270**: 580-585.

Kyte, J. and Doolittle, R. F. (1982) A simple method of displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.

Lahue, E.E. and Matson, S. W. (1990) Purified *E.coli* F-factor TraY protein binds *oriT*. *J. Bacteriol.* **172**: 1385-1391.

Matson, S.W. and Morton, B. S. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J. Biol. Chem.* **266**: 16232-16237.

Provencher, S.W. and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**: 33-37.

Reygers, U., Wessel, R. Müller, H. and Hoffmann-Berling, H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J.* **10**: 2689-2694.

Schwab, M., Gruber, H. and Hogenauer, G. (1991) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5**: 439-446.

Strickland, E.H. (1974) Aromatic contribution to circular dichroism spectra of proteins. *CRC Crit. Rev. Biochem.* **2:** 113-175

Thompson, R. and Taylor, L. (1982) Promoter Mapping and DNA Sequencing of the F Plasmid Transfer Genes *traM* and *traJ*. *Mol. Gen. Genet.* **188**: 513-518.

Willetts, N. and Maule, J. (1985) Specificities of IncF plasmid conjugation genes. *Genet. Res.* **47**: 1-11.

Willetts, N. and Skurray, R. (1987) Structure and Function of the F Factor and Mechanism of Conjugation, p. 1110-1133. In F.C. Neidhardt, J.L. Ingraham, B. Magasanik, K.B. Low, M. Schaechter and H.E. Umbarger, (ed.) *Escherichia coli* and *Salmonella typhimurium* Cellular and Molecular Biology. American Society For Microbiology, Washington.

Willetts, N. and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48**: 24-41.

CHAPTER VI

The Interaction of Integration Host Factor with the Origin of Transfer of the IncFV Plasmid pED208[1]

A. Introduction

Integration host factor (IHF) is a sequence specific DNA-binding protein from *E. coli* that plays a major role in many bacterial processes such as bacteriophage λ DNA integration into the bacterial chromosome (Kim *et al.*, 1990), viral packaging (Saha *et al.*, 1990, Shinder and Gold, 1989), transposon recombination (Gamas *et al.*, 1987b), control of expression of bacterial and viral genes (Giladi *et al.*, 1990; Mahajna *et al.*, 1986; Pereira *et al.*, 1988; Tsui *et al.*, 1991), and plasmid replication (Gamas *et al.*, 1986). More recently, IHF binding sites were identified in the control region of nitrogen fixation operons of a number of purple bacteria (Hoover *et al.*, 1990). IHF belongs to a class of prokaryotic DNA-binding proteins called "histone-like" because of their ability to fold the DNA into nucleosome-like structures (Drlica and Rouviere-Yaniv, 1987; Friedman, 1988). In particular, IHF recognizes specific sequences in AT-rich regions and, upon binding, bends the DNA by more than 140° (Kosturko *et al.*, 1989; Prentki *et al.*, 1987). In bacterial conjugation, plasmid DNA is nicked at a specific site called *oriT* and the nicked strand is transferred from the donor cell into the recipient cell in a 5'-3' direction (Thompson *et al.*, 1984). The region of DNA containing the nick site and the sites of action of the various gene products involved in the transfer process is called the *oriT* region. IHF$^-$ strains show reduced transfer efficiency of the conjugative plasmids F and R100 (Dempsey 1987; Gamas *et al.*, 1987a), and several binding sites for IHF have been identified and characterized in the *oriT* regions of these plasmids (Dempsey, and Fee, 1990; Inamoto *et al.*, 1990).

---

[1] A version of this chapter has been submitted for publication. Di Laurenzio, L., Frost, L.S., Scraba, D.G., Paranchych, W. J. Bacteriol. 1992.

pED208 is the derepressed form of the IncFV plasmid F$_0$*lac* (Finlay *et al.*, 1986). Sequence analysis of the region responsible for the conjugative transfer of DNA of pED208 has revealed extensive similarities at the level of gene organization and protein sequence with the transfer region of F, the paradigm of conjugative plasmids, although the two plasmids share little actual DNA sequence homology (Finlay *et al.*, 1984).

In this study, we have investigated the interaction between IHF and the *oriT* region of pED208. We demonstrate that the organization of these binding sites is similar to those of the other IncF plasmids.

B. Experimental Procedures

1. Plasmids

All the clones used in this work were derived from the vector pUC18 (Yanisch-Perron *et al.*, 1985) and are shown in Fig. VI. 1.

2. Recombinant DNA techniques

Restriction enzymes and the large fragment of DNA polymerase I (Klenow) were purchased from Boehringer-Mannheim (Laval, Quebec, Canada) or Bethesda Research Laboratories (Gaithersburg, MD). The enzymes were used according to the supplier's specifications.

3. Fragment labeling

Plasmids pLDL101, pLDL102 (Chapter II) and pLDL107 (Fig. VI. 1) were digested with *Eco*RI and *Hin*dIII or *Pst*I to obtain different fragments of the *oriT* region (for specific enzymes see legends of each figure). The *Eco*RI site is at the beginning of the *oriT* region (Chapter II), whereas the *Hin*dIII and *Pst*I sites are in the multiple cloning site of the vector pUC18 (Yanisch-Peron *et al.*, 1985). The products of digestions were separated on 1.6% agarose gels, and the DNA fragments containing the *oriT* region were purified using the Elutrap electro-elution system (Schleicher & Schuell, Keene, NH). The purified fragments were then specifically labeled either at the 3' end by the "filling-in" reaction with the large fragment of DNA polymerase (Klenow) or at the 5' end by the T4 kinase labeling reaction (Maniatis *et al.*, 1982).

4. Electrophoretic mobility retardation assay

DNA-protein complexes were detected using the method described by Garner and Revzin (1981) with the following modifications: incubations of labeled DNA fragments with protein were performed in a 20 µl volume for 20 min at room temperature in a "DNase I footprinting buffer" modified for IHF binding (25 mM HEPES, pH 7.8, 0.5 mM dithiothreitol (DTT), 0.05 mM disodium EDTA, 5%

glycerol, 0.5 mM phenylmethylsulphonyl fluoride (PMSF)) containing either 100 mM KCl or 100 mM monopotassium glutamate and 2 mg ml^{-1} bovine serum albumin (BSA) (Sigma Chemicals, St. Louis, MO). Reaction mixtures were loaded on pre-electrophoresed 5% polyacrylamide non-denaturing gels without the addition of loading buffer (Kosturko *et al.*, 1989) and electrophoresis was performed at 30 mA at room temperature. The dried gels were exposed to X-ray film (Eastman Kodak Co, Rochester, NY or Fuji, Tokyo, Japan) at -70°C.

Quantitative measurements of the bands on the autoradiograms were performed using a Chromoscan 3 densitometer (Joyce Loebl, Gateshead, England).

5. Calculation of the binding constants

The molar ratio of protein to DNA that caused band retardation in the electrophoretic mobility retardation assays was between $5 \cdot 10^4$ and $5 \cdot 10^5$ to 1; thus the effective protein concentration is not decreased by subtracting the fraction of protein bound to the DNA. Under these conditions, the thermodynamic binding constant (K_b) can be estimated using the following formula (Prentki *et al.*, 1987):

$$K_b = \frac{F_{bound}}{[P]\,(1 - F_{bound})}$$

where F_{bound} is the fraction of DNA bound, as determined by densitometric analysis of the autoradiograms and [P] is the total concentration of protein in the reaction mixture, assuming that the protein is 100% active. Thus the reported values may be an underestimation of the real K_b.

6. DNase I protection analysis

DNase I (Deoxyribonuclease I, Worthington Biochemicals, Freehold, NJ) protection analyses were performed on labeled DNA fragments as previously described (Chapter II) using the DNase I footprinting buffer described above. Fifty $\mu g\ ml^{-1}$ of poly(dI-dC)poly(dI-dC) (Sigma Chemicals) was added as competitor DNA. The products of the reaction were separated on 6% or 12% sequencing gels. Radioactivity was detected by exposure of the dried gels to X-ray film at -70°C.

7. Electron microscopy

The electron microscopy characterization of the DNA-protein complexes was performed in collaboration with Dr. D. G. Scraba. The carbon replicas were prepared by Roger Bradley. In order to visualize the complexes between the pED208 *oriT* region and pure IHF, a replica method of preparation (Lurz *et al.*, 1986) was used. DNA at a concentration of 1 $\mu g\ ml^{-1}$ was incubated with 2 $\mu g\ ml^{-1}$ of pure IHF in a 5 µl reaction volume for 15 min at room temperature. The reaction mixture was then diluted six times in binding buffer (25 mM HEPES, pH 7.8, 0.5 mM DTT, 0.05 mM EDTA, 1mM $MgCl_2$) containing either 100 mM KCl or 100 mM monopotassium glutamate, and allowed to adsorb for 1 min to a 1 cm square of freshly cleaved mica, which was placed on top of a 25 µl droplet of the solution. Excess liquid was removed with a filter paper, after which the square was floated on a 2% aqueous solution of uranyl acetate and on 3 successive droplets of water. After blotting on filter paper, the mica square was placed on a rotating stage in a Balzers BA511-M metal shadowing apparatus equipped with electron guns for evaporation and a quartz-crystal film thickness monitor. The DNA on the mica square was shadowed, while rotating, with 1.8 nm of platinum at an angle of 8°, and a surface replica was made by evaporation of 15 nm of carbon from above.

The carbon replica was floated off the mica onto a water surface and picked up on a 300 mesh copper grid covered with a film of parlodion. Electron microscopy was carried out with a Philips EM420-STEM system. The operating voltage was 100 kV, the C1 aperture was 50 μm, and bright field/dark field images were recorded from a high resolution photomonitor on 35 mm Pan-X film (Kodak).

Contour length measurements of the DNA molecules were made using a Hewlett-Packard digitizer (Model 9874A) connected to Tetronix 4051 computer.

C. Results

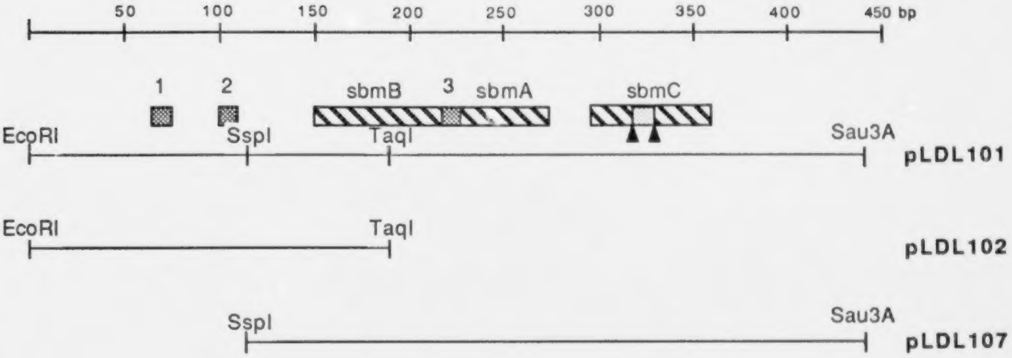
1. Electrophoretic mobility retardation assays of the *oriT* region incubated with IHF

Comparison of the sequence of the *oriT* region (Chapter II) with the consensus sequence for IHF binding (Gamas *et al.*, 1987a) identified three potential sites of interaction in this region (Fig. VI. 1, top, **1**, **2** and **3**). These sites contained one non-identical base and were located at positions 65-77, 100-109 and 218-227, respectively (Fig. VI. 1, bottom).

Electrophoretic mobility retardation assays were used to study the binding of pure IHF (kindly donated by Dr. H. Nash, NIH) to a ^{32}P-labeled DNA fragment containing the *oriT* region. The ^{32}P-end-labeled *Eco*RI/*Pst*I DNA fragment from

Figure VI. 1

Linear map of the fragments used in this Chapter. The position of potential IHF binding sites and known TraM binding sites is shown. Plasmids pLDL101 and pLDL102 have been described in Chapter II; pLDL107 was obtained by subcloning the *SspI/Hin*dIII fragment from pLDL101 into the *SmaI/Hin*dIII sites of the vector pUC18. The stippled boxes numbered **1**, **2** and **3** represent the probable IHF binding sites identified by sequence comparison to a consensus sequence reported by Gamas *et al.* (1987b). The striped boxes labeled *sbmA*, *sbmB* and *sbmC* represent the specific binding sites of the TraM protein, respectively. The dotted box within *sbmC* is a region of partial protection and the arrowheads are the hypersensitive sites as discussed in Chapter II. In the lower portion of this figure the 5' to 3' nucleotide sequence of the three potential binding sites are compared to the consensus sequence for IHF binding. The position of the three sequences is reported and is based on the sequence presented in Chapter II. The non-conserved nucleotides are indicated with an x.


50
100
150
200
250
300
350
400
450 bp
1
2
sbmB
3
sbmA
sbmC
EcoRI
SspI
TaqI
Sau3A
pLDL101
EcoRI
TaqI
pLDL102
SspI
Sau3A
pLDL107

IHF consensus $\begin{array}{l}\mathrm{TAANNNNTTG}\\ \mathrm{C}\end{array}$

site 1 $^{75}\mathrm{T}\overset{x}{\mathrm{T}}\mathrm{AAATCTTG}^{66}$

site 2 $^{109}\mathrm{TAATACA}\overset{x}{\mathrm{C}}\mathrm{TG}^{100}$

site 3 $^{227}\mathrm{TAATTTA}\overset{x}{\mathrm{G}}\mathrm{TG}^{218}$

pLDL101, containing the 440 bp *Eco*RI/*Sau*3A fragment from the *oriT* region (Fig. VI 1), was incubated with IHF in a buffer containing 0.1 M KCl, and the complexes were separated on 5% non-denaturing polyacrylamide gels (Fig. VI. 2). Two bands of different mobility, compared to the free fragment, were identified (Fig. VI. 2. Complex I and II). Complex II was sensitive to the addition of increasing amounts of calf thymus DNA as competitor, the only DNA-protein complex detectable at high concentrations of calf thymus DNA was complex I (data not shown). This suggested a specific interaction of IHF with the DNA site in complex I. The amounts of IHF used in the electrophoretic mobility retardation assays ranged between 5 fmol and 50 pmol, corresponding to concentrations between 0.25 nM and 2.5 μM; these are comparable to the estimated intracellular concentration of IHF (1.5 μM) (Kosturko *et al.*, 1989); thus the observed interactions could occur *in vivo*. The binding constants for IHF binding to the high and low affinity sites were calculated to be equal to or greater than $2.1 \cdot 10^{7}\ M^{-1}$ and $2.1 \cdot 10^{5}\ M^{-1}$, respectively. These values are comparable to those calculated by Tsai *et al.* (1990) for the IHF sites in the F system, and to the IHF binding sites in the IS1 element (Prentki *et al.*, 1987).

2. DNase I protection analysis of the *oriT* region incubated with IHF

Mapping of the sites of interaction of IHF with the *oriT* region was achieved by DNase I footprint analysis of the *Eco*RI/*Hin*dIII fragment of pLDL102 containing the 189 bp *Eco*RI/*Taq*I fragment from the *oriT* region (Fig. VI. 1) incubated with IHF. This fragment contains only sites 1 and 2 of the three potential IHF binding sites. The result of the protection analysis is shown in Figure VI. 3. The region between nucleotides 80 and 115, containing site 2, was protected from DNase I

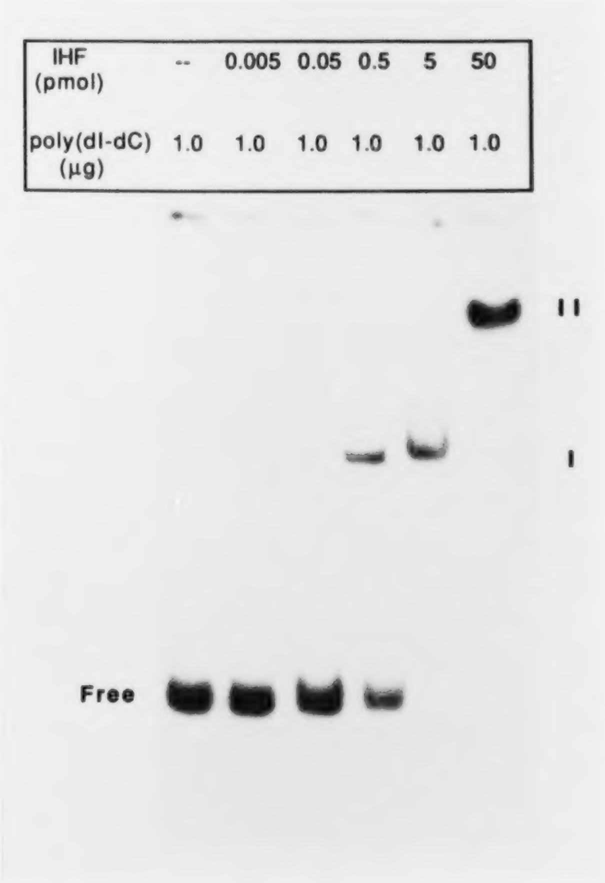
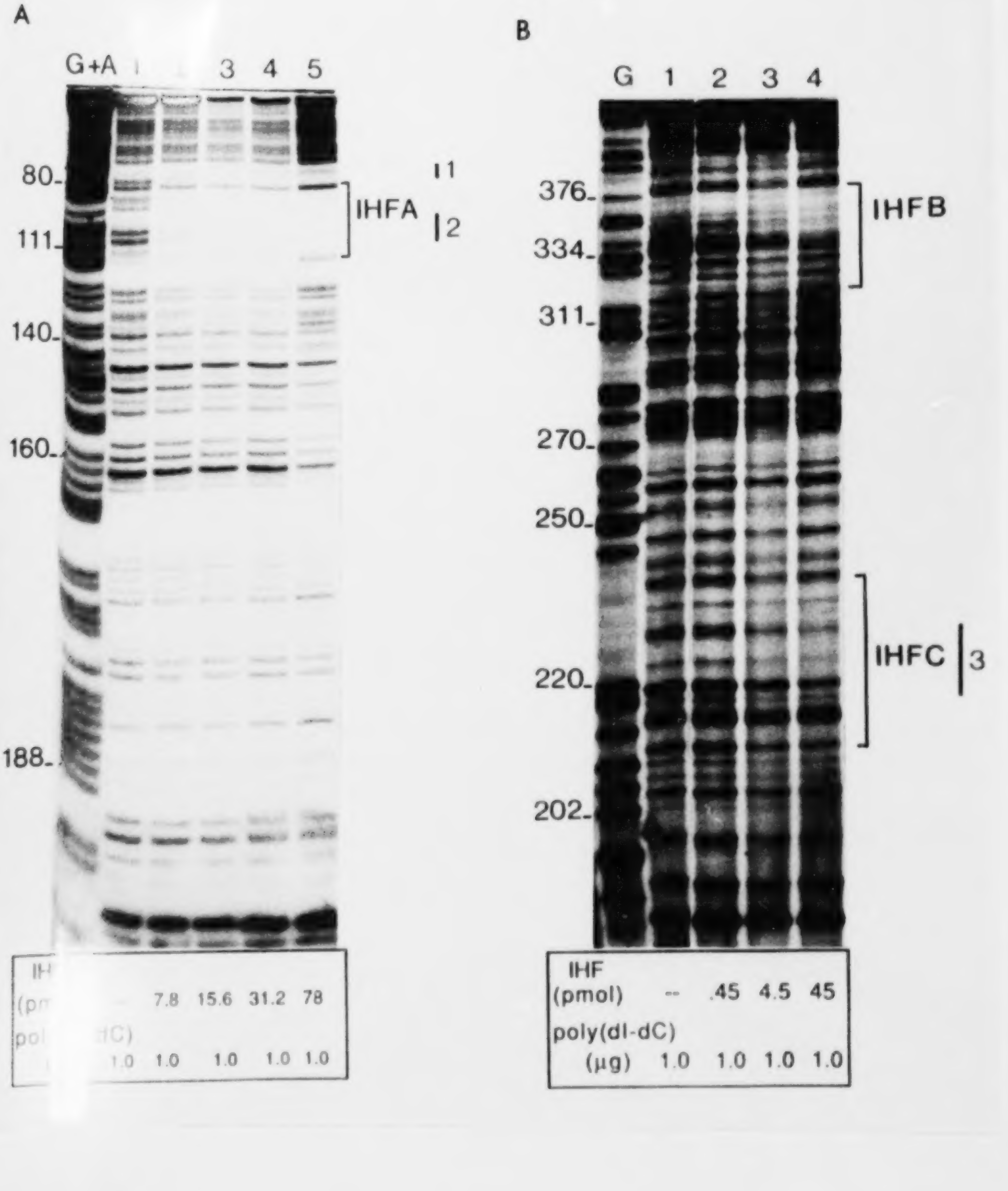

IHF (pmol)
--
0.005
0.05
0.5
5
50
poly(dI-dC) (µg)
1.0
1.0
1.0
1.0
1.0
1.0
II
I
Free

Figure VI. 2

Binding of IHF to the *oriT* region of pED208. 0.1 fmol of the ^{32}P-labeled *Eco*RI/*Pst*I fragment from pLDL101 was incubated with increasing concentrations of pure IHF. The mobility of free DNA and of the DNA-protein complexes I and II are indicated.

Figure VI. 3

DNase I protection of pED208 *oriT* region incubated with IHF. **A**. DNase I footprinting analysis of 9.5 fmol of the 189 bp *Eco*RI/*Taq*I fragment of the *oriT* region (pLDL102) incubated with IHF. The *Eco*RI/*Hin*dIII fragment from pLDL102 was selectively labeled at the *Hin*dIII site (upper strand). Lane G+A; Maxam and Gilbert sequencing reaction. **B**. DNase I footprinting analysis of 25 fmol of the 350 bp *Eco*RI/*Pst*I fragment (pLDL107) incubated with IHF in the presence of potassium glutamate. The *Eco*RI/*Pst*I fragment containing the *Ssp*I/*Sau*3A fragment of the *oriT* region was labeled at the *Eco*RI site (lower strand). Lane G, Maxam and Gilbert sequencing reaction. The numbers at the left-hand of both panels represent the position in the sequence from the *Eco*RI site in the *oriT* region (Chapter II). The protected regions are shown in brackets. The concentrations of IHF and poly(dI-dC)poly(dI-dC) are given for each lane in the box at the bottom of the figure. The thick lines represent the potential binding sites **1**, **2** and **3**.


A
G+A
1
2
3
4
5
80
111
140
160
188
IHFA
1
2
IH
(pm
--
7.8
15.6
31.2
78
pol
dC)
1.0
1.0
1.0
1.0
1.0
B
G
1
2
3
4
376
334
311
270
250
220
202
IHFB
IHFC
3
IHF
(pmol)
--
.45
4.5
45
poly(dI-dC)
(µg)
1.0
1.0
1.0
1.0

digestion and was named IHFA (Fig. VI. 3 **A**). Site **1** was not protected. In order to test site **3**, DNase I footprinting was performed on the *Eco*RI/*Hin*dIII DNA fragment from plasmid pLDL101, containing the 440 bp *Eco*RI/*Sau*3A fragment and all three potential binding sites, incubated with IHF. Only one protected region corresponding to IHFA was identified (data not shown). Thus IHFA corresponded to the high affinity site identified by the electrophoretic mobility retardation assay.

In order to increase the affinity of IHF for other possible binding sites in the *oriT* region, DNase I protection analysis was performed on a fragment of the *oriT* region which did not include site A in a reaction buffer containing 0.1 M monopotassium glutamate instead of potassium chloride (Richey *et al.*, 1987) and 50 $\mu g\ ml^{-1}$ of poly(dI-dC)poly(dI-dC) as competitor DNA. The 365 bp *Eco*RI/*Pst*I fragment from pLDL107 containing the 325 bp *Ssp*I/*Sau*3A fragment from nucleotide 117 to 440 of the *oriT* region (Fig. VI. 1) was incubated with IHF and digested with DNase I. The products of digestions were separated on a 6% polyacrylamide sequencing gel and are shown in Figure 3 **B**. Two protected regions were identified; the first one between nucleotides 210 and 240 (IHFC) contains site **3**; the second footprint, between nucleotides 315 and 375 (IHFB) had not been predicted to contain a potential binding site. Instead it covers the second half of the third binding site of the TraM protein (Fig. VI. 1, sbmC).

From electrophoretic mobility retardation assays performed in the presence of monopotassium glutamate and poly(dI-dC)poly(dI-dC), estimates of the binding constants (Tsai *et al.*, 1990) for the two sites were determined to be $2.4 \cdot 10^7\ M^{-1}$ for IHFB and greater than $2.5 \cdot 10^6\ M^{-1}$ for IHFC, respectively. The difference of two orders of magnitude between the values of K_b for IHFB in the presence or absence of monopotassium glutamate is reminiscent of the effect of glutamate on the interaction of RNA polymerase with promoter sites (Leirmo *et al.*, 1987).

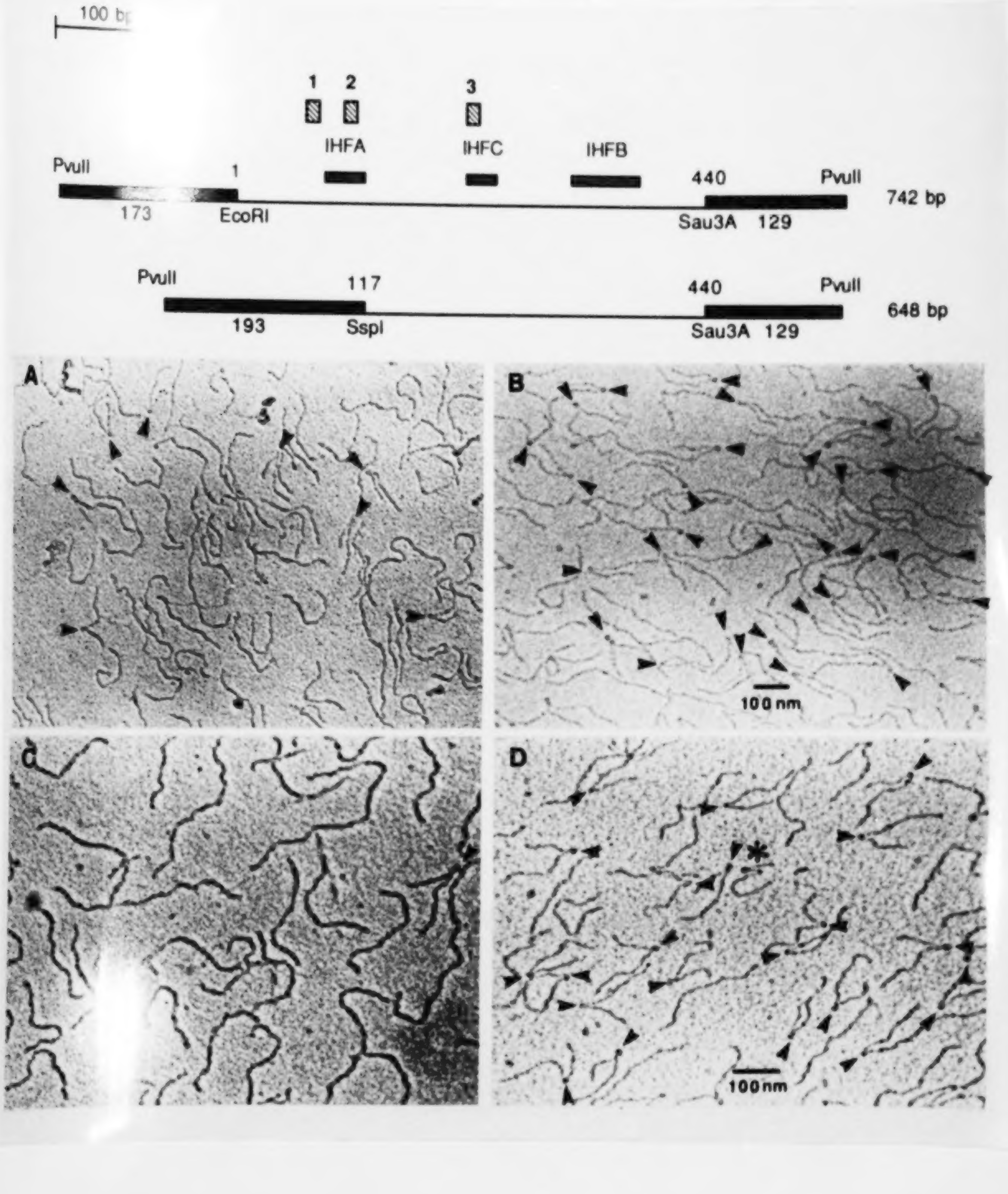
3. Electron microscopy characterization of the IHF-*oriT* complexes

IHF has been shown to cause a sharp bend in the DNA, which is visible in the electron microscope (Kosturko *et al.*, 1989). In order to visualize the complexes between the *oriT* region and IHF, electron microscopy of carbon replicas of two fragments of the *oriT* region (Fig. VI. 4, top) incubated with IHF was performed. The electron micrographs showed DNA molecules that were bent at an angle of more than 140° with a protein molecule bound to the apex. When the 742 bp *Pvu*II DNA fragment from pLDL101, containing the 440 bp *Eco*RI/*Sau*3A fragment from the *oriT* region, was incubated with IHF, 14% of the molecules (89 out of 632) were bent (Fig. VI. 4, **A**). When the incubation buffer contained 0.1 M in monopotassium glutamate, the percentage rose to 59% (181 out of 307) (Fig. VI. 4, **B**). The 648 bp *Pvu*II DNA fragment from pLDL107, containing the *Ssp*I/*Sau*3A fragment from the *oriT* region (nucleotides 117 to 440), which contains site **3** and is missing site A (nucleotides 80 to 115), showed 1.7% (6 out of 357) bent molecules in the absence of monopotassium glutamate (Fig. VI. 4, **C**); this increased to 59% (246 out of 419) in its presence (Fig. VI. 4, **D**).

In order to localize the sites of binding of IHF, contour lengths of a number of molecules from representative fields were measured. The position of the bend in 58 of the 742 bp long molecules corresponded to nucleotide 103 (± 17 bp) in the *oriT* map in the absence of monopotassium glutamate. Measurements of 88 of the

Figure VI. 4

Electron microscopy of IHF-*oriT* region complexes. Top. Linear map of the DNA fragments used in electron microscopy. The 742 bp and 648 bp fragments were obtained by digestion of pLDL101 and pLDL107, respectively, with *Pvu*II and the fragments containing the *oriT* region were purified as described in Experimental Procedures. The thick bars represent vector DNA (pUC18) and their lengths are indicated below the line, the numbers above the map line represent the position in the pED208 sequence (Chapter II). Bottom. Electron micrographs of representative fields of IHF incubated with the two DNA fragments in the absence or presence of potassium glutamate. **A** and **B**, the 742 bp fragment was incubated with IHF as described in Experimental Procedures in the absence or presence of monopotassium glutamate in the reaction buffer, respectively. **C** and **D**, the 648 bp fragment incubated with IHF in the absence or presence of potassium glutamate in the reaction buffer, respectively. The arrows indicate bent DNA molecules with an electron-dense spot (the IHF protein) at the apex. The asterisk indicates a DNA molecule with two bends.


100 bp
1
2
3
IHFA
IHFC
IHFB
PvuII
1
440
PvuII
742 bp
173
EcoRI
Sau3A
129
PvuII
117
440
PvuII
648 bp
193
SspI
Sau3A
129
A
B
100 nm
C
D
*
100 nm

molecules incubated with IHF in the presence of monopotassium glutamate gave similar values for the position of the bend (99 ±19 bp). Thus the position of the bend on the 742 bp long molecule, determined in the presence or absence of monopotassium glutamate, corresponded to IHFA. When 78 of the 648 bp long DNA molecules were measured, the position of the bend was found to correspond to nucleotide 330 (±15 bp) on the *oriT* region map, which was in agreement with the binding site for IHFB. A few molecules that showed two bends were identified. Measurements of the contour lengths of the arms of the 742 bp (5 molecules) or 648 bp (11 molecules) long DNA fragments with two bends showed a random distribution of the second bend and its position was not assigned. The histogram of the results of the contour measurements are shown in Figure VI. 5.

4. The interactions of IHF and TraM with the *oriT* region are not co-operative

It has been shown that the transfer protein TraM binds to three sites in the *oriT* region of pED208 (Chapter II). Since IHFB and IHFC are adjacent to, or overlap the TraM binding sites (Fig. VI. 5), the effect of incubating both IHF and TraM with the *oriT* region was investigated by means of DNase I footprinting. The *Eco*RI/*Pst*I fragment from pLDL102 containing the 189 bp fragment of the *oriT* region incubated with either IHF or TraM or both proteins together, was briefly digested with DNase I and the products of digestion separated on a 12% polyacrylamide sequencing gel. IHF was found to bind to the high affinity site A, independently from TraM binding to *sbmB* at position 150-189 (Fig. VI. 1), since the interaction of one protein with the DNA was not affected by the presence of the other (data not shown).

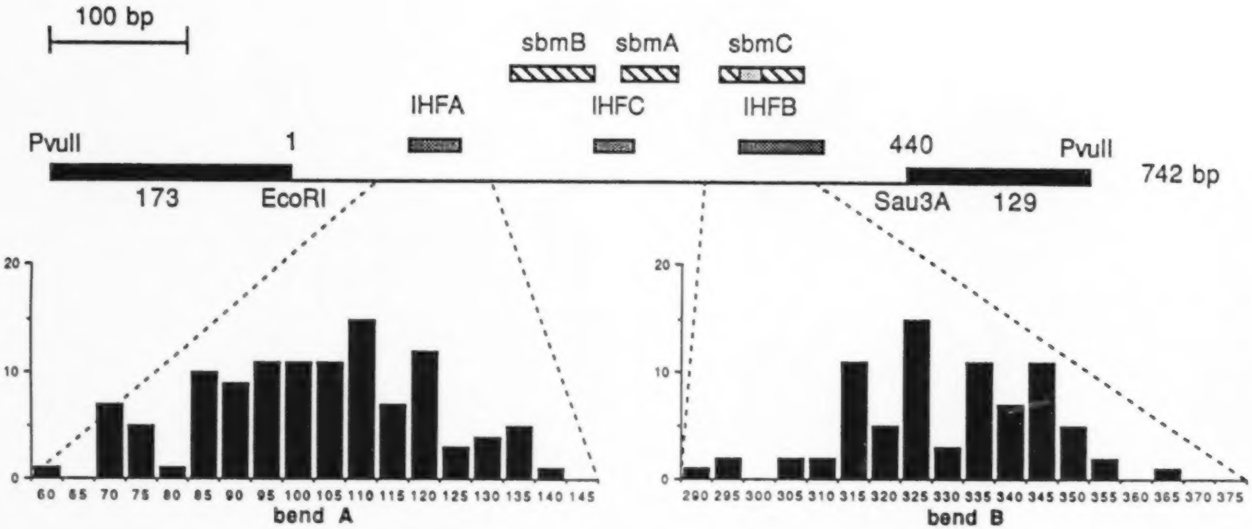


Figure VI. 5

Position of the bends in the 742 bp and 648 bp fragments as observed by electron microscopy. The top line represents the map of the *oriT* region as shown in Figure VI. 4. The TraM binding sites are represented as striped boxes and the IHF footprints are represented by stippled boxes. The histograms show the distribution of the location of the bends in the 742 bp fragment and in the 648 bp fragment obtained by measuring 136 and 78 molecules, respectively.

A similar set of experiments was performed using the *Eco*RI/*Hin*dIII fragment from pLDL101 containing the 440 bp *Eco*RI/*Sau*3A fragment incubated with both proteins in a reaction buffer containing 0.1 M monopotassium glutamate. Although the two low affinity IHF binding sites (IHFB and IHFC) partially overlapped the TraM binding sites *sbmA* and *sbmC*, no co-operative effect of binding was detected (Fig. VI. 6). A short region within IHFB was not protected by IHF binding; this divided the footprint into two smaller footprints; b' spanning between nucleotide 315 and 340, and b" between nucleotide 345 and 375. IHFB was not identified as a possible IHF binding site by computer analysis of the sequence and, although it contains a stretch of A and T residues (nucleotides 317 to 330, Chapter II), it has very little homology to the consensus sequence. However, IHFB was protected by lower concentrations of IHF protein than IHFC, which contains site 2 of high homology to the IHF binding consensus sequence (Fig. VI. 1).

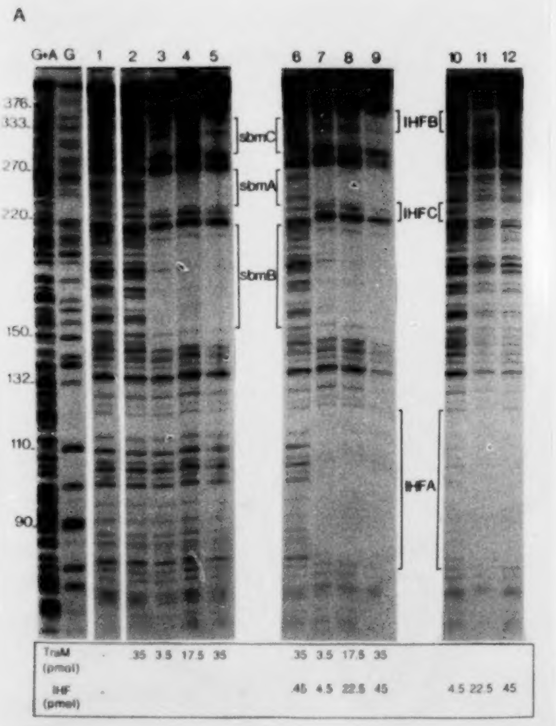
D. Discussion

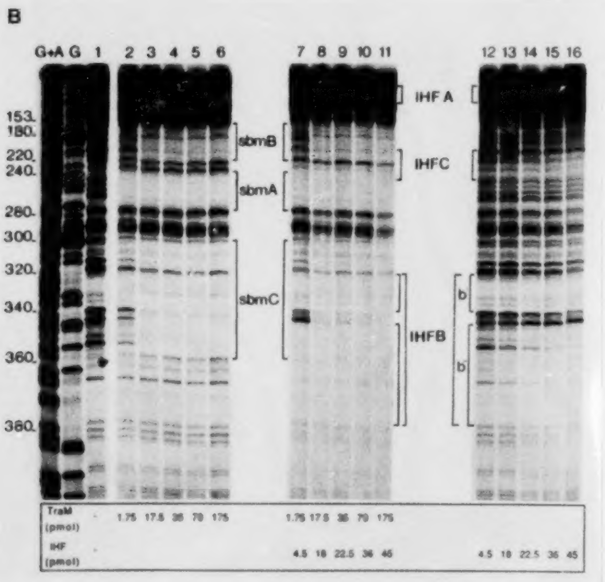
This study demonstrate the interaction of IHF protein with the origin of transfer of the F-like plasmid pED208. Electrophoretic mobility retardation assays and DNase I protection analysis identify three sites of interaction; binding of IHF to two of them could be visualized by electron microscopy as causing sharp bends in the DNA.

The IHFA site is located between the putative site of specific nicking, based on sequence comparison with the *oriT* regions of other F-like plasmids, and the first binding site of the TraM protein (Chapter II). The location of this site is consistent

Figure VI. 6

DNase I protection of pED208 *oriT* region incubated with IHF and TraM. **A**. DNase I protection analysis of 20.0 fmol of the *Eco*RI/*Pst*I fragment from pLDL101 labeled at the *Eco*RI site (bottom strand). **B**. DNase I protection analysis of 16 fmol of the *Kpn*I/*Hin*dIII fragment from pLDL101 labeled at the *Hin*dIII site (upper strand). Lane 1, free DNA. Lane G+A and Lane G, Maxam and Gilbert sequencing reactions. The amounts of TraM and IHF in each reaction are given in the box below the footprints. The numbers at the left hand side represent the position in the sequence from the *Eco*RI site in the *oriT* region. The footprinted regions are shown in brackets. **b'** and **b"** indicate the smaller footprints discussed in the Results section.


A
G+A G 1 2 3 4 5 6 7 8 9 10 11 12
376.
333.
270.
220.
150.
132.
110.
90.
sbmC
sbmA
sbmB
IHFB
IHFC
IHFA
TraM (pmol) - .35 3.5 17.5 35 35 3.5 17.5 35
IHF (pmol) - .45 4.5 22.5 45 4.5 22.5 45


B
G+A G 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
153.
180.
220.
240.
260.
300.
320.
340.
360.
360.
sbmB
sbmA
sbmC
IHF A
IHFC
IHFB
b
b'
TraM (pmol) - 1.75 17.5 35 70 175 1.75 17.5 35 70 175
IHF (pmol) - 4.5 18 22.5 36 45 4.5 18 22.5 36 45

with a similar site found in F and R100 plasmids (Tsai *et al.*, 1990; Dempsey and Fee, 1990) (Fig. VI. 7), where the region between IHFA and the first TraM binding site is the site of action of the TraY protein (Inamoto and Ohtsubo, 1990; Lahue and Matson, 1990), which has been postulated to be part of the nicking complex together with the TraI protein (Everett and Willetts, 1980). IHF binding to site A may bring into close proximity the nick site and the TraYI complex bound to the DNA as proposed by Inamoto *et al.* (1990), thus favouring efficient nicking and initiation of DNA transfer.

Unlike many DNA-binding proteins, IHF decreases local negative supercoiling by adding positive supercoils to the DNA (Snyder *et al.*, 1989). If IHF binding occurs in a region of the DNA that is a closed topological domain, the resulting torsional strain is energetically unfavourable, and the neighbouring DNA will undergo local untwisting in order to resolve the unstable configuration. This model has been proposed for the mechanism of action of the intasome at the *attP* site of phage λ and it suggests a dual function for IHF: to bring into close proximity the integrase binding sites and to promote untwisting of the DNA at *attP*, thus initiating recombination.

It has recently been shown in the F system that negative supercoiling affects the efficiency of nicking of the TraI protein at *oriT* (Matson and Morton, 1991; Reygers *et al.*, 1991). In conjugative DNA transfer, IHF binding to the *oriT* region may create a metastable topological configuration that would be resolved by local untwisting of the DNA to form a region of single stranded DNA (Travers, 1990);

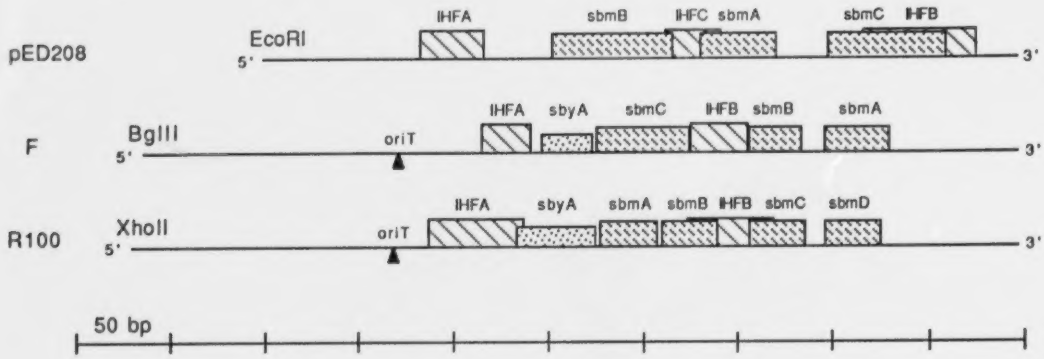

pED208
EcoRI
5'
3'
IHFA
sbmB
IHFC
sbmA
sbmC
IHFB
F
BglII
oriT
IHFA
sbyA
sbmC
IHFB
sbmB
sbmA
R100
XhoII
oriT
IHFA
sbyA
sbmA
sbmB
IHFB
sbmC
sbmD
50 bp

Figure VI. 7

Diagram of the interactions at the *oriT* region of the IncF plasmids pED208, F and R100. The data on pED208 are based on this as well as Chapter II, the data on the F *oriT* region are based on published work (Reygers *et al.*, 1991; Lahue and Matson, 1990, Tsai *et al.*, 1990) as well as Chapter IV. The data on the R100 *oriT* region are taken from Inamoto and Ohtsubo (1990), Dempsey and Fee (1990), Inamoto *et al.* (1990), Abo *et al.* (1991), Inamoto *et al.* (1991). The 5' end of the F *oriT* region is the *Bgl*II site at 66.7 kbp on the map of F (Willetts and Skurray, 1980), for the R100 *oriT* region is the *Xho*II site (Inamoto *et al.*, 1991) and for the pED208 *oriT* is the *Eco*RI site (Chapter II). The 3' end of all the *oriT* regions corresponds to the ATG start codon of *traM.*

this would then be a suitable substrate for nicking and initiation of DNA transfer. Interestingly, the sequence between IHF and the TraM binding sites is AT-rich and is the binding site for the TraY protein. This would also support the idea that the role of IHF is two-fold; to bend the DNA and to promote unwinding of sequences next to its binding site. The role of TraY could then be seen as maintaining this underwound sequence.

In the pED208 system, the traY gene product as well as the *oriT* site have not been characterized, although a region encoding an open reading frame of high homology to the TraY protein of F and R100 has been identified (Finlay *et al.*, 1986). The region between IHFA and *sbmB* of pED208 *oriT* region could accommodate binding of the TraY protein in analogy to what is seen in the F and R100 systems (Inamoto and Ohtsubo, 1990, Lahue and Matson, 1990) (Fig. VI. 7). Comparison of the sequences of the *oriT* regions of several F-like plasmids suggests that pED208 *oriT* nick site may be located around nucleotide 60 (Chapter II).

Since the sequence periodicity of the TraM binding motif is also greater than 10.5 bp (11-12 bp) (Chapter II), the TraM protein may, like IHF, add positive supercoils (Travers, 1990) and further increase the unwinding of the neighbouring DNA.

Electron microscopy of purified TraM protein bound to the *oriT* region was also performed (data not shown). The TraM protein, when present predominantly as tetramers, did not bend the DNA. However, at higher concentrations of proteins, aggregates of TraM were visible, which appeared to shorten the length of the DNA molecules significantly (50 bp/ aggregate) upon binding to the DNA. The significance of the compaction of the DNA by TraM is not known, but resembles the activity of the TraK protein in the RP4 relaxosome (Ziegelin, 1989).

The other two sites of interaction of IHF with the *oriT* region could be localized only in the presence of monopotassium glutamate in the reaction mixture. This salt is present in cells at a physiological concentration of between ~0.03 and ~0.25 molal and is known to increase the affinity of DNA-binding proteins for their target DNA (Leirmo *et al.*, 1987; Richey *et al.*, 1987). The function of these sites is not clear at this moment. We propose that IHFC, which is located between two TraM binding sites, might facilitate the interaction of TraM molecules bound at the two different sites or provide two binding sites for the same TraM molecule, which is a tetramer of the same subunit (Chapter V).

A fraction of the DNA fragments containing the *oriT* region of pED208 showed abnormal mobility in non-denaturing gel electrophoresis. The percentage of DNA molecules with decreased electrophoretic mobility ranged between 5% and 50%, as the monopotassium glutamate concentration increased from zero to 0.2 M in the reaction buffer (data not shown). Decreased electrophoretic mobility is a typical feature of bent DNA (Hagerman, 1990) and it has been recently shown by Tsai *et al.* (1990) that the *oriT* region of F has two bends. Bending of the *oriT* region may be a common feature among the F-like plasmids and may play a role in the transfer process. However, the number and the location of the bend(s) in the pED208 *oriT* region is not known.

In the F and R100 systems (Tsai *et al.*, 1990; Dempsey and Fee, 1990) a second IHF binding site has been identified in the *oriT* region; this site has low affinity for IHF and is located between two TraM binding sites (Fig. VI. 7), corresponding to the site with the lowest affinity (IHFC) for the IHF protein in the pED208 *oriT* region.

It was shown in Chapter II of this thesis that *sbmC* contains a region (nucleotide 310 to 330) which is partially protected from DNase I digestion and is delimited by hypersensitive sites (Fig. VI. 1). This region is protected by IHF and corresponds to IHFB. The importance of this region of overlap between TraM and IHF binding sites is not clear; however IHF binding to IHFB may play a role in the regulation of the expression of the *traM* gene.

DNase I protection analysis of the *oriT* region incubated in the presence of both IHF and TraM did not show any co-operative effect between the proteins, suggesting independent interaction with linear DNA. It is possible that IHF could have a higher affinity for IHFB and IHFC in supercoiled DNA and could affect *traM* gene expression by changing the local level of supercoiling. We are currently investigating the effects of IHF on the expression of *traM* and other *tra* genes *in vivo*. While IHF$^-$ strains have been shown to have decreased transfer efficiency of F and R100 (Dempsey, 1987; Gamas *et al.*, 1987a), similar experiments were not possible with pED208 since the plasmid was unstable in the IHF$^-$ strains described by Gamas *et al.* (1987a).

E. Bibliography

Abo, T., Inamoto, S. and Ohtsubo, E. (1991) Specific DNA Binding of the TraM Protein to the *oriT* Region of Plasmid R100. *J. Bacteriol.* **173**: 6347-6354.

Dempsey, W.B. (1987) Integration host factor and conjugative transfer of the antibiotic resistance plasmid R100. *J. Bacteriol.* **169**: 4391-4392.

Dempsey, W.B. and Fee, B. E. (1990) Integration host factor affects expression of two genes at the conjugal transfer origin of plasmid R100. *Mol. Microbiol.* **4**: 1019-1028.

Drlica, K. and Rouviere-Yaniv, J.(1987) Histonelike proteins of bacteria. *Microbiol. Rev.* **51**: 301-319.

Everett, R. and Willetts, N. (1980) Characterisation of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136**: 129-150.

Finlay, B.B., Frost, L. S. and Paranchych, W. (1984) Localization, cloning, and sequence determination of the conjugative plasmid ColB2 pilin gene. *J. Bacteriol.* **160**: 402-407.

Finlay, B.B., Frost, L. S. and Paranchych, W. (1986) Nucleotide sequence of the *traYALE* region from IncFV plasmid pED208. *J. Bacteriol.* **168**: 990-998.

Friedman, D.I. (1988) Integration host factor: a protein for all reasons. *Cell* **55**: 545-554.

Gamas, P., Burger, A-C., Churchward, G., Caro, L., Galas, D. and Chandler, M. (1986) Replication of pSC101: effects of mutations in the *E. coli* DNA binding protein IHF. *Mol. Gen. Genet.* **204**: 85-89.

Gamas, P., Caro, L., Galas, D. and Chandler, M. (1987) Expression of F transfer functions depends on the *E.coli* integration host factor. *Mol. Gen. Genet.* **207**: 302-305.

Gamas, P., Chandler, M. G., Prentki, P. and Galas, D. J. (1987) *Escherichia coli* Integration host factor binds specifically to the ends of the insertion sequence *IS1* and to its major insertion hot-spot in pBR322. *J. Mol. Biol* **195**: 261-272.

Garner, M.M. and Revzin, A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucleic Acids Res.* **9**: 3047-3060.

Giladi, H., Gottesman, M. and Oppenheim, A. B. (1990) Integration host factor stimulates the phage lambda pL promoter. *J. Mol. Biol.* **213**: 109-121.

Hagerman, P. J. (1990) Sequence-directed curvature of DNA. *Annu. Rev. Biochem.* **59**: 755-781.

Hoover, T.R., Santero, E., Porter, S. and Kustu, S. (1990) The integration host factor stimulates interaction of RNA polymerase with NIFA, the transcriptional activator for nitrogen fixation operons. *Cell* **63**: 11-22.

Inamoto, S., Abo, T.and Ohtsubo, E.(1990) Binding sites of integration host factor in *oriT* of plasmid R100. *J. Gen. Appl. Microbiol.* **36**: 287-293.

Inamoto, S. and Ohtsubo, E. (1990) Specific binding of the TraY protein to *oriT* and the promoter region for the *traY* gene of plasmid R100. *J. Biol. Chem.* **265**: 6461-6466.

Inamoto, S., Yoshioka, Y. and Ohtsubo, E. (1991) Site- and strand-specific nicking *in vitro* at *oriT* by the TraY-TraI endonuclease of plasmid R100. *J. Biol. Chem.* **266**: 10086-10092.

Kim, S., Moitoso de Vargas, L., Nunes-Duby, S. E. and Landy, A. (1990) Mapping of a higher order protein-DNA complex: two kinds of long-range interactions in λ *att*L. *Cell* **63**: 773-781.

Kosturko, L. D., Daub, E. and Murialdo, H. (1989) The interaction of *E.coli* integration host factor and lambda cos DNA: multiple complex formation and protein-induced bending. *Nucleic Acids Res.* **17**: 317-334.

Lahue, E. E. and Matson, S. W. (1990) Purified *E. coli* F-factor TraY protein binds *oriT*. *J. Bacteriol.* **172**: 1385-1391.

Leirmo, S., Harrison, C., Cayley, C. D., Burgess, R. R. and Record, T.M. Jr. (1987) Replacement of potassium chloride by potassium glutamate dramatically enhances protein-DNA interactions *in vitro*. *Biochemistry* **26**: 2095-2101.

Lurz, R., Grote, M., Dijk, J., Reinhardt, R. and Dobrinski, B. (1986) Electron microscopic study of DNA complexes with proteins from the Archebacterium *Sulfolobus acidocaldarius*. *EMBO J.* **5**: 3715-3721.

Mahajna, J., Oppenheim, A. B. Rattray, A. and Gottesman, M. (1986) Translation initiation of bacteriophage lambda gene cII requires integration host factor. *J Bacteriol.* **165**: 167-174.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning. A laboratory manual, Cold Spring Harbor Laboratory Press, New York.

Matson, S.W. and Morton, B. S. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J. Biol. Chem.* **266**: 16232-16237.

Pereira, R.F., Ortuno, M. J. and Lawther, R. P. (1988) Binding of integration host factor (IHF) to the *ilvGp1* promoter of the *ilvGMEDA* operon of *Escherichia coli* K12. *Nucleic Acids Res.* **16**: 5973-5989.

Prentki, P., Chandler, M. and Galas, D. J. (1987) *Escherichia coli* integration host factor bends the DNA at the ends of *IS1* and in an insertion hotspot with multiple IHF binding sites. *EMBO J.* **6**: 2479-2487.

Reygers, U., Wessel, R., Müller, H. and Hoffmann-Berling, H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J.* **10**: 2689-2694.

Richey, B., Cayley, S. C., Mossing, M. C., Kolka, C., Anderson, C. F., Farrar, T. C. and Record, T. Jr. (1987) Variability of the intracellular ionic environment of *Escherichia coli*. *J. Biol. Chem.* **262**: 7157-7164.

Saha, S., Haggard-Ljungquist,E. and Nordstrom, K.(1990) Integration host factor is necessary for lysogenization of *Escherichia coli* by bacteriophage P2. *Mol. Microbiol.* **4**: 3-11.

Shinder, G. and Gold, M. (1989) Integration host factor (IHF) stimulates binding of the gpNu1 subunit of λ terminase to *cos* DNA. *Nucleic Acids Res.* **17**: 2005-2022.

Snyder, U.K., Thompson, J. F. and Landy, A. (1989) Phasing of protein-induced DNA bends in a recombination complex. *Nature* **341**: 255-257.

Thompson, R., Taylor, L., Kelly, K.,Everett, R. and Willetts, N. (1984) The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks. *EMBO J.* **3**: 1175-1180.

Travers, A.A. (1990) Why bend DNA? *Cell* **60**: 177-180.

Tsai, M-M., Fu Y. H. F.and Deonier, R.C. (1990) Intrinsic bends and integration host factor binding at F plasmid *oriT*. *J. Bacteriol.* **172**: 4603-4609.

Tsui, P., Huang, L. and Freundlich, M. (1991) Integration host factor binds specifically to multiple sites in the *ompB* promoter of *Escherichia coli* and inhibits transcription. *J.Bacteriol.* **173**: 5800-5807.

Willetts, N. and Skurray, R. (1980) The conjugation system of F-like plasmids. *Annu. Rev. Genet.* **14**: 41-76.

Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

Ziegelin, G. (1989) Konjugativer transfer des IncPα plasmids RP4: die spezifitätsdeterminanten codieren *oriT*-bindungsproteine. Dissertation, Freie Universität Berlin, Germany.

CHAPTER VII

Final Conclusions

Ten years ago TraM was believed to be encoded by an open reading frame adjacent to the *oriT* region (Thompson and Taylor, 1982). Because F *traM*$^-$ mutants produce pili and λ*oriT*$^+$ transducing phages from these cells have nicked genomes, it was thought that the TraM function was associated with the actual process of DNA transfer (Everett and Willetts, 1980; Kingsman and Willetts, 1978). In 1987 there were rumors that F TraM had been extensively characterized and publication of those results was imminent, thus my project dealt with the *traM* gene of the F-like plasmid pED208 with a view to comparing the two systems. The data on F TraM were never published, and in 1989 after a meeting on conjugation organized by Dr. Karin Ippen-Ihler at Texas A&M, I cloned the F *traM* gene and started the characterization of the F gene product as well as pursuing the purification of pED208 TraM.

In this Chapter I would like to address some of the questions that remain unanswered and suggest future research.

A. pED208 Plasmid.

Very little is known about this plasmid. Most of the work was performed in Dr. Paranchych's laboratory and dealt with the characterization of the pilus, the pilin subunit and the definition of an epitope (Worobec *et al.*, 1985). Only a handful of *tra* genes have been sequenced and they show similarities to the *tra* genes of F (Finlay *et* *al.*, 1983; 1986a; 1986b; Chapter II) suggesting a common mechanism of conjugation for F and pED208.

The characterization of the factors interacting at the *oriT* region of pED208, when compared with the system of other F-like plasmids, will help elucidate the relationship between the transfer operon and the *oriT* region. The *traY* operon contains the genes necessary for pilus biosynthesis and assembly and some of the genes involved in DNA transfer; in particular, the first gene of the operon, *traY*, is interesting because of its postulated dual function: (1) as a component of the nicking complex (Willetts and Wilkins, 1984), and (2) as one of the regulators of the expression of the transfer operon, (Inamoto *et al.*, 1990). Therefore *traY* gene expression may co-ordinate pilus biosynthesis with conjugal DNA metabolism.

In pED208, transfer operon expression is derepressed as a result of the insertion of an IS2 element at the beginning of the *traY* gene. It is presently unclear whether pED208 produces a truncated TraY protein, missing the first 13 amino acids, or whether it does not have a functional *traY* gene product. If the latter hypothesis is true, how does the nicking complex of pED208 work? Is TraY protein necessary? One of the advantages of pED208 over F is that the former has a naturally repressed counterpart F_o*lac*, which has been shown to be identical to pED208 in the *traY* region except for the IS2 element (Chapter II).

Until the genetic map of the *tra* region of pED208 is known and a collection of *tra* mutants is obtained, it will be difficult to elucidate the pED208 TraM function in conjugation, since TraM is plasmid-specific. Nevertheless, the biochemical characterization of pED208 TraM may provide us with useful information on the structure and functional domains of the protein.

The amino acid sequence alignment of the five homologs of TraM (Fig. VII. 1) shows a high degree of sequence identity among the proteins, with most of the differences occurring at the amino terminal end. This suggests that the amino terminus comprises the DNA binding domain, since the known TraM proteins recognize different sequences (Abo *et al.*, 1991; Schwab *et al.*, 1991a; Chapter II and IV). Deletion mutations and site-specific mutations will aid in identification the amino acids involved in DNA recognition, in multimerization, and in signalling to start DNA transfer.

Two aspects of the *in vitro* characterization of the interactions of TraM with the *oriT* region are puzzling. First, it is not clear why it is necessary to add a vast excess of TraM to the *oriT* region to detect a shift in DNA mobility in gel electrophoresis, and second, why the DNA-protein complexes appear as broad bands.

TraM may interact with the *oriT* region without causing a change in the mobility of the DNA (under the conditions employed), until the concentration of TraM in the reaction tube is such that it forms aggregates, which may promote DNA looping, resulting in an all-or-none electrophoretic mobility effect. This electrophoretic behavior is typical of proteins that bind DNA in a co-operative fashion. Looped DNA structures were observed with the electron microscope when concentrated samples of TraM were used in binding studies with the *oriT* region

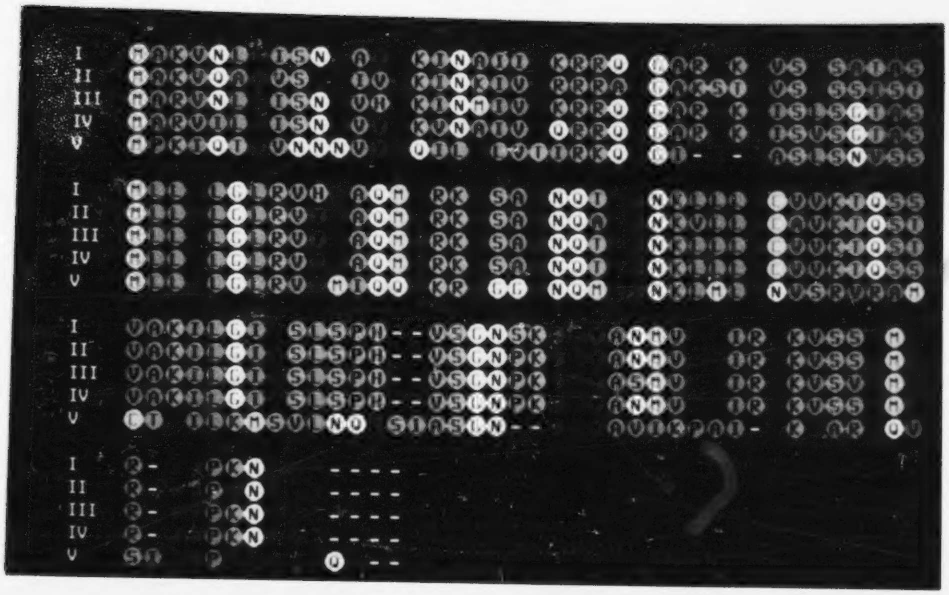


Figure VII. 1

Amino acid sequence alignment of the five TraM prototypes. The amino acids are reported in one-letter code and colour-coded according to the nature of the side chain (blue for basic residues, red for acidic residues, green for hydrophobic, brown for aromatic, yellow for polar, *etc.*). The five sequences were aligned by Dr. Laura Frost and Brett Finlay and they represent the TraM protein of the following plasmids: I=F, II=ColB4, III=R1, IV=R100 and V=pED208. The software for this Figure was developed by David Bacon in Dr. Wayne Anderson's laboratory (Department of Biochemistry).

(discussed in Chapter VI). Two additional pieces of evidence support the co-operative binding of TraM molecules to the *oriT* region: (i) retardation assays of aliquots taken from DNaseI protection reactions do not show retardation on native gel electrophoresis, but do show footprints on sequencing gels (data not shown); and (ii) electrophoretic mobility retardation assays of IHF incubated with the *oriT* region, performed under the same conditions as TraM binding assays, showed discrete retarded bands (Chapter VI).

Crystals of pure TraM have been grown and, although of acceptable size, they diffract poorly (Dr. R. Read, personal communication). The reasons for this are not known, but gel exclusion chromatography of the purified protein sample confirmed a homogeneous population of tetramers (A. Boodhoo, personal communication).

The *oriT* region of pED208 contains a large number (17) of similar pentameric sequences evenly spaced within the TraM binding sites (Chapter II). Studies using synthetic oligonucleotides showed TraM bound to pairs of these sequences which were separated by one turn of the DNA helix (Chapter III). The minimal binding sequence recognized by TraM should provide a suitable substrate for the crystallographic analysis of the interactions between the amino acid residues involved in DNA binding. Before attempting to co-crystallize TraM with double-stranded oligonucleotides, further characterization of the sequence-specificity of the synthetic DNA-substrate is required. Competition assays of TraM binding to the double-stranded template **T-T/A-A** using template **T-A** duplex as competitor and retardation assays of TraM binding to oligonucleotides with sequences identical to *sbmA* should determine which is the optimal binding site.

Moreover, the strength and stochiometry of TraM binding to the oligonucleotides needs to be elucidated. The electrophoretic mobility retardation assays shown in Chapter III did not clarify the number of DNA binding sites present per TraM tetramer, because the labeled DNA was in limiting amounts. Electrophoretic mobility retardation assays of the oligonucleotide-TraM complexes in which the TraM concentration is kept constant, while the amount of labeled DNA is increased, would determine the molar ratio of tetramer protein to DNA, at which the TraM binding sites are saturated. In addition, the rate of TraM binding to the oligonucleotides was not determined. Thus, it is possible that TraM interaction with the oligonucleotides is a rapid exchange between bound and unbound forms, whereas the interaction with the *oriT* region may have a longer half-life with smaller binding constants. Binding assays performed over a period of several hours will allow the estimation of the exchange rate constants.

Finally, regulation of the expression of the *traM* gene is an issue barely touched by the results presented in this thesis, but nevertheless a very important aspect of the regulation of the whole conjugative process. The promoter region of pED208 *traM* has not been mapped and it is not known if one (like in F, Thompson and Taylor, 1982) or more transcripts (like in R1 and R100, Koronakis *et al.*, 1985; Dempsey, 1989) are produced.

B. F Plasmid

The F plasmid is considered to be the paradigm of all conjugative plasmids (Willetts and Skurray, 1987). Unlike pED208, a wide number of mutants have been created and among them there is a *traM*- mutant (F*lac traM*102). Very little is known about this mutant except that it is not a nonsense mutation and it does not affect ColE1 mobilization (Van de Pol *et al.*, 1978; Willetts, 1980). Characterization of the phenotype of the mutant may help understanding the function of TraM.

A minimal binding site for F TraM could not be identified easily in the F *oriT* region in analogy to pED208 TraM. F TraM may interact with its binding sites differently than pED208 TraM does. In the F *oriT* region only two pairs of the ten GGPyGC sequences are 12 bp apart and they are found in the higher affinity sites *sbmA* and *sbmB* located in the *traM* promoter region (see Chapter IV, Fig. IV. 7). The DNase I protection pattern of the *oriT* region differs between the two plasmids. pED208 TraM binds equally to both DNA strands whereas F TraM appears to prefer one strand; the order of binding to the *sbm* sites is also different.

Both proteins tend to form aggregates at high concentration as shown by the sedimentation equilibrium studies. This feature is shared with R1 TraM (Schwab *et al.*, 1991a). The intracellular concentration of TraM is not known; thus it is unclear if aggregation occurs *in vivo* or is an artefact of the purification procedure. Sedimentation velocity analysis of TraM would identify the critical concentration that causes aggregation and comparison with the intracellular concentration of TraM would clarify this aspect .

In studies on the characterization of the interactions of the F TraM protein with the F *oriT* region, the protein was found to bind to a sequence in the mobilizable plasmid ColE1 *oriT* region, which is homologous to the *sbmC* site in the F *oriT* region (Willetts and Wilkins, 1984; Thompson *et al.*, 1984) (Chapter IV). ColE1 plasmid encodes the genes necessary for the nicking reaction and plasmid-DNA mobilization (*mob* genes) but it cannot transfer unless a conjugative plasmid is present (Warren et al., 1979). Furthermore, ColE1 can be mobilized by a limited range of conjugative plasmids (Reeves and Willetts, 1974), implying that other functions, beside the establishment of a "tunnel" between the mating cells, are required for triggering plasmid mobilization (Willetts and Wilkins, 1984). One of these functions could be the putative mating signal protein TraM. The results presented in Chapter IV suggest the involvement of TraM protein in the DNA transfer of ColE1 plasmid.

The immunoblots of cell fractions may be a very sensitive assay to elucidate the function of TraM in DNA transfer. The amount of TraM found associated with the membrane fraction increased in the presence of the *oriT* region and the F plasmid (discussed in Chapter IV). Thus TraM localization in strains carrying mutations in different *tra* genes may identify the factor(s) responsible for anchoring TraM to the inner membrane. The genes that I would be most interested in assaying would be *traY* and *traD*, followed by *traG* and *traN*. The lack of a *traY* mutant (Willetts and Wilkins, 1984) suggested either lack of a phenotype or a lethal mutation. The effect of a null mutation in *traY* has not been adequately determined.

One of the limitations of the *in vitro* experiments presented in this thesis is that the characterization of the protein binding was performed on linear DNA fragments. Recent data on TraI (Reygers *et al.*, 1991; Matson and Morton, 1991) showed that supercoiling is necessary for TraI activity; it is therefore possible that the local supercoiling of the *oriT* region influences TraM interaction with this region or *vice versa. In vivo* "footprinting" of donor cells or mating cells may answer this question.

In vitro studies aimed at reconstituting the multiprotein-DNA complex acting at *oriT* would help to elucidate the role played by each protein and the protein-protein interactions in analogy to what has been shown in the promiscuous plasmid RP4 (Pansegrau *et al.*, 1990b).

During conjugation mediated by the promiscuous plasmid RP4, specific transfer gene products interact with the *oriT* region and assemble a protein-DNA complex called the plasmid DNA-protein relaxation complex or "relaxosome". All the factors involved in relaxosome formation are encoded by the "relaxase operon" which is 3' from the nick site, as is the transfer region of F-like plasmids (Ziegelin *et al.*, 1991). Four RP4 proteins have been implicated to be involved in relaxosome assembly: TraI, TraH, TraJ and TraK (Ziegelin *et al.*, 1989; Furste *et al.*, 1989; Pansegrau *et al.*, 1990b; Ziegelin *et al.*, 1991). These proteins have been purified and used for *in vitro* relaxosome-reconstitution experiments (Pansegrau *et al.*, 1990a).

The proposed model of relaxosome assembly involves the following steps. The TraJ protein binds to an imperfect 19 bp long inverted repeat (Ziegelin *et al.*, 1989); this step is essential for binding of the next protein, TraI, to the *oriT* region. TraI possesses potential site- and strand-specific endonuclease activity. Upon treatment of RP4 relaxosomes with protein denaturing agents, open circular species are isolated that have a covalently bound TraI protein at the 5' end of the nicked DNA strand (Pansegrau *et al.*, 1990b). The *oriT* site was mapped one turn of the DNA helix away from the TraJ protein binding site, suggesting interactions between TraJ and TraI (Pansegrau *et al.*, 1990b). The interactions between TraI and TraJ are stabilized by the action of the acidic protein TraH, which does not have DNA binding activity and has been shown by centrifugation on glycerol gradients to form stable complexes with TraI and TraJ in the absence of DNA (Pansegrau *et al.*, 1991). The TraK protein which has been shown to be essential for DNA transfer *in vivo*, may also play a role in relaxosome assembly. TraK binds to a 200 bp long sequence in the *oriT* region of RP4 and wraps the DNA into nucleosome-like structures around a central core of protein as shown by electron microscopy (Ziegelin *et al.*, 1991). It has been proposed that this structure may induce topological changes in its vicinity, which would increase the yield of relaxed intermediates (Pansegrau *et al.*, 1991).

The similarities between TraK of RP4 and TraM of F and pED208 plasmids are striking. The proteins exhibit anomalous behavior in electrophoretic mobility retardation assays; they show an all-or-none DNA binding activity in gel electrophoresis and DNA-protein complexes appear in the presence of large excesses of protein over DNA. The binding region of TraK was identified on an intrinsically curved DNA segment and the binding sites were found to be spaced by one turn of the DNA helix (Ziegelin, *et al.*, 1991). TraK appears to have multiple functions in RP4-mediated DNA transfer. It is required to obtain transconjugants, *traK*$^-$ mutants show a ≈200-fold reduction in transfer efficiency, and although it is not necessary for the nicking reaction, TraK increases the efficiency of the reaction both *in vitro* and *in vivo*. It is not known if TraK has a signalling function similar to TraM. Amino acid sequence comparison between F TraM and TraK of RP4 do not show sequence homology (Dr W. Pansegrau, personal communication).

C. The Icarus Model

On the basis of the data presented in this thesis and published by other groups, plus many long conversations with Dr. Laura ("Daedalus") Frost, I would like to present a model for the conjugal transfer of DNA. This model is called Icarus because, like the mythological hero, it is trying to fly with wings that are not very strong and may not survive the heat of the Sun.

In non-mating donor cells, TraM concentration builds to a level where the protein binds to *sbmA* and *sbmB*, thus repressing its own expression (Schwab *et al.*, 1991b). It is partitioned between the cytoplasm and the inner membrane, although the site of binding in the membrane is not known (Chapter II and IV). Meanwhile TraYI in the membrane nicks and ligates *oriT* in a reversible manner (Everett and Willetts, 1980), or it nicks and is bound to the ends like type I topoisomerase-DNA complexes but without the strand exchange. At any given time there is an equilibrium between the two states represented in Fig. VII. 2 (**A** and **B**).

Once the mating pair is formed and stabilized, the small amount of TraM in the membrane is modified/released and binds to *sbmC*, separating TraI from TraY (Fig. VII. 2, **C**), or TraM binding to *sbmC* forms a nucleosome-like structure (analogous to TraK in RP4) which would change the topology of the *oriT* region locally (Fig.

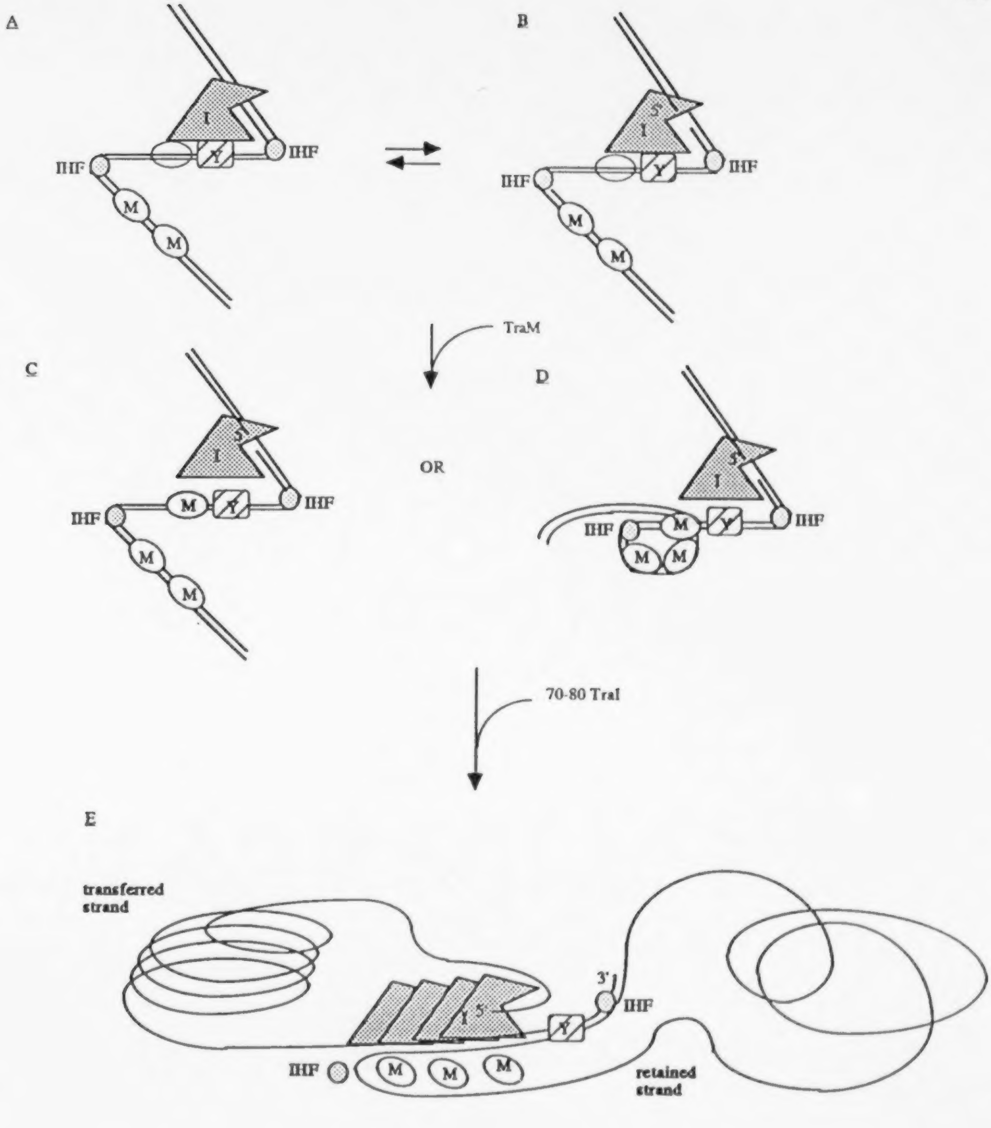

A
B
C
D
E
I
Y
M
IHF
5'
3'
TraM
OR
70-80 TraI
transferred strand
retained strand

Fig. VII. 2

The Icarus model. A schematic diagram of the model discussed in the text is presented. For simplicity the diagram does not indicate the mating pair junction or the complementary DNA strands. The plasmid-encoded proteins TraI (I), TraY (Y) and TraM (M) as well as the host factor IHF are indicated.

VII. 2, **D**) and might prolong the life time of the nicked species. TraM binding to *sbmC* results in the inhibition of the ligase activity in TraI; thus the TraI molecule is converted to a covalently bound form to the 5' end of the nicked strand (Willetts and Wilkins, 1984), and aggregates with 70 to 80 TraI molecules to make the helicase complex (Abdel-Monem *et al.*, 1977). The TraI agglomerate unwinds the nicked strand and presents it to TraD at the junction between the mating cells (Willetts and Wilkins, 1984).

Once most of the nicked strand has been transferred into the recipient cell, the helicase unwinds the TraM binding sites and displaces TraM from the *sbmC* site thus re-activating the ligase activity of the TraYI complex and the 5' end is ligated to the 3' end before leaving the donor cell (Fig. VII. 2, **E**). This mechanism resembles stage II of the replication of bacteriophage ϕX174. In this system, the supercoiled double-stranded form of the phage genome is used as template for the synthesis of new viral genome copies. The viral gpA protein initiates the replication process by introducing an endonucleolytic cleavage on one of the strands at the origin of replication. gpA binds covalently to the 5' end of the cleaved DNA strand of the replicative form II (RFII). The *E. coli rep* protein then binds to gpA of the RFII-gpA complex and unwinds the double-stranded DNA. The 5' end which is covalently attached to gpA travels along the replication fork in a looped rolling-circle, whereas the 3' end is used by the DNA polymerase III holoenzyme as primer for the synthesis of the complementary strand. When one round of the looped rolling-circle synthesis of the viral strand is completed, the gpA bound at the 5' end of the parental viral strand cuts the viral strand at the newly synthesized origin and rejoins the two parental strands to generate circular viral DNA (Arai *et al.*, 1981).

TraM is expressed and binds to the newly replicated double-stranded *sbmA* and *sbmB* sites as well as to the unknown receptor in the membrane of the donor cell waiting for the next mating pair to be formed.

Integration Host Factor plays a dual function in this model: first, by bending the DNA between the TraY binding site and the *oriT* site, it brings the two sites into close proximity, thus stabilizing the interactions between the TraYI nicking complex and the *oriT* site, and second, it may prevent the free 3' end of the nicked strand from unwinding.

The mechanism of TraM binding to the *oriT* region may be analogous to the interactions of phage λ repressor with its operators. The repressor is a dimer of a polypeptide of 27,000 Da. The amino terminus contains the DNA binding domain whereas the carboxy terminal domain is responsible for multimerization. The repressor regulates the expression from two divergent promoters P_R and P_L, which control the "genetic switch" between the lytic and lysogenic states of λ phage. The right and left operators (O_R and O_L) each contains three binding sites for the repressor, which binds to the first two sites of each operator co-operatively. Co-operative binding to the two sites is the result of the interactions between the carboxy terminal domains of repressor molecules bound to these sites. The repressor has higher affinity for the first site of each operator compared to the second site, and binding to the latter is increased by the presence of repressor bound to the first site. Co-operativity does not take place when the repressor binds the third site, because the distances between the second and the third sites do not allow inter-chain contacts between the carboxy termini of repressor molecules bound at these sites. At concentrations usually found in a lysogen, two sites are occupied in each operator region but not the third one. The balance between lysogeny and the lytic cycle depends on the cellular concentration of repressor. Intact repressor is present in the cell at a concentration sufficient to ensure that the first two sites of each operator are occupied. Under these conditions, the amino terminus of one of the repressor subunits bound to O_{R2} interacts with RNA polymerase and promotes expression of its own gene. When the intracellular concentration of repressor is high enough, the repressor binds to the third site O_{R3} and prevents RNA polymerase from binding to the repressor maintenance promoter (P_{RM}) promoter, which overlaps the right promoter, thus turning off the expression of the repressor gene. Like the λ repressor, TraM binds two sites (*sbmA*, *sbmB*) with higher affinity compared to the third site (*sbmC*). It is not known if binding to *sbmA* and *sbmB* is a co-operative event and it would be interesting to create mutations of these sites and analyze the interactions of TraM with the mutated sited in analogy to what has been done in the λ system. Unlike the λ repressor, which activates its own gene expression, TraM binding to the first two sites may prevent RNA polymerase from binding to the *traM* promoter, thus maintaining a limiting amount of protein in the cell. Binding to the third site would be "the switch" to start DNA transfer.

The basic questions "What is the function of TraM? Does TraM signal that transfer can start? How does it signal that transfer can start ?" remain unanswered. The "link" between the establishment of the mating pair and the initiation of DNA transfer is presently unknown. During the first steps of conjugation (Recognition, Attachment, Retraction and Stabilization, Fig. I. 2) the conjugative pilus plays an essential role. It is possible that the pilus itself or one of the factors involved in pilus assembly acts as the signal: for example, TraM or its "anchor" to the membrane may sense the accumulation of pilin subunits in the membrane as a result of the retraction event, causing the release/modification of TraM from the membrane and consequent binding to the *oriT* region.

The similarities between the TraI, TraJ and TraK proteins of the promiscuous plasmid RP4 and the F-encoded TraI, TraY, and TraM, respectively, suggest a common mechanism of conjugal DNA transfer. Furthermore, the tumor-inducing plasmids and the mobilizable plasmids have mechanisms of DNA mobilization that are similar to the conjugative plasmids, suggesting that the wide spectra of gene transfer mechanisms observed in nature may have evolved from a common ancestor.

"The greatest part of happiness is wisdom;
.....and wisdom we learn only when we are old"

Sophocles, Antigone

D. Bibliography

Abdel-Monem, M., Lauppe, H.-F., Kartenbeck, J., Dürwald, H.,Hoffman-Berling, H. (1977) Enzymatic unwinding of DNA: mode of action of *Escherichia coli* DNA unwinding enzyme. *J. Mol. Biol* **110**: 667-685.

Abo, T., Inamoto, S. and Ohtsubo, E. (1991) Specific DNA Binding of the TraM Protein to the *oriT* Region of Plasmid R100. *J. Bacteriol.* **173**: 6347-6354.

Arai, N., Polder, L., Arai, K., Kornberg, A. (1981) Replication of ϕX174 DNA with purified enzymes. II. Multiplication of the duplex form by coupling of continuous and discontinuous synthetic pathways. *J. Biol. Chem.* **256**: 5239-5247.

Dempsey, W.B. (1989) Sense and antisense transcripts of *traM*, a conjugal transfer gene of the antibiotic resistance plasmid R100. *Mol. Microbiol.* **3**: 561-570.

Everett, R. and Willetts, N. (1980) Characterisation of an in vivo system for nicking at the origin of conjugal DNA transfer of the sex factor F. *J. Mol. Biol.* **136**: 129-150.

Finlay, B.B., Frost, L.S. and Paranchych, W. (1986a) Nucleotide Sequence of the *traYALE* Region from IncFV Plasmid pED208. *J. Bacteriol.* **168**: 990-998.

Finlay, B.B. and Paranchych, W. (1986b) Nucleotide sequence of the surface exclusion genes *traS* and *traT* from the IncF$_0$*lac* plasmid pED208. *J. Bacteriol.* **166**:713-721.

Finlay, B.B., Paranchych, W. and Falkow, S. (1983) Characterization of conjugative plasmid EDP208. *J. Bacteriol.* **156**: 230-235.

Fürste, J.P., Pansegrau, W., Ziegelin, G., Kroger, M., Lanka, E. (1989) Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. U.S.A.* **86**: 1771-1775.

Inamoto, S., Ohtsubo, E. (1990) Specific binding of the TraY protein to *oriT* and the promoter region for the *traY* gene of plasmid R100. *J. Biol. Chem.* **265**: 6461-6466.

Kingsman, A. and Willetts, N. (1978) The requirements for conjugal DNA synthesis in the donor strain during f*lac* transfer. *J. Mol. Biol.* **122**: 287-300.

Koronakis, V.E., Bauer, E. and Hogenauer, G. (1985) The *traM* gene of the resistance plasmid R1: comparison with the corresponding sequence of the *Escherichia coli* F factor. *Gene* **36**: 79-86.

Matson, S.W. and Morton, B.S. (1991) *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. *J. Biol. Chem.* **266:** 16232-16237.

Pansegrau, W., Balzer, D., Kruft, V., Lurz, R., Lanka, E. (1990a) *In vitro* assembly of relaxosomes at the transfer origin of plasmid RP4. *Proc. Natl. Acad. Sci. U.S.A.* **87**: 6555-6559.

Pansegrau, W., Ziegelin, G. and Lanka, E. (1990b) Covalent association of the *traI* gene product of plasmid RP4 with the 5'-terminal nucleotide at the relaxation nick site. *J. Biol. Chem.* **265**: 10637-10644.

Pansegrau, W., Ziegelin, G. and Lanka, E. (1991) Enzymology of bacterial conjugation: *in vitro* reconstitution of relaxosomes of plasmid RP4. EMBO Workshop on bacterial conjugative systems. pp 88-91.

Reeves, P. and Willetts, N. (1974) Plasmid specificity of the origin of transfer of sex factor F. *J. Bacteriol.* **120**: 125-130.

Reygers, U., Wessel, R., Müller, H. and Hoffman-Berling, H. (1991) Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. *EMBO J.* 10: 2689-2694.

Schwab, M., Gruber, H. and Hogenauer, G. (1991a) The TraM protein of plasmid R1 is a DNA-binding protein. *Mol. Microbiol.* **5**: 439-446.

Schwab, M. and Hogenauer, G. (1991b) Studies on the expression of TraM of plasmid R1. EMBO Workshop on bacterial conjugation systems. p 69.

Thompson, R. and Taylor, L. (1982) Promoter Mapping and DNA Sequencing of the F Plasmid Transfer Genes *traM* and *traJ*. *Mol. Gen. Genet.* **188**: 513-518.

Thompson, R., Taylor, L., Kelly, K., Everett, R.,Willetts, N. (1984) The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks. *EMBO J.* **3**: 1175-1180.

Van de Pol, H., Veltkamp, E. and Nijkamp, H.J.J. (1978) Genetic analysis of the mobilization of the non-conjugative plasmid CloDF13. *Mol. Gen. Genet.* **160:** 139-149.

Warren, G.J., Saul, M.W. and Sherratt, D.J. (1979) ColE1 plasmid mobility: essential and conditional functions. *Mol. Gen. Genet.* **170:** 103-107.

Willetts, N. (1980) Interactions between the F conjugal transfer system and CloDF13::TnA plasmids. *Mol Gen. Genet.* **180:** 213-217.

Willetts, N. and Wilkins, B. (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48**: 24-41.

Willetts, N.S. and Skurray, R. (1987) Structure and function of the F factor and mechanism of conjugation, p. 1110-1133. In F.C. Neidhardt, J.B. Ingraham, K.B. Low, B. Magasanik, M. Schaecter and H.E. Umbarger, (ed.) *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular biology. American Society for Microbiology, Washington, D.C..

Worobec, E.A., Paranchych, W., Parker, J.M., Taneja, A.K. and Hodges, R.S. (1985) Antigen-antibody interaction. The immunodominant region of EDP208 pili. *J. Biol. Chem.* **260:** 938-943.

Ziegelin, G., Fürste, J.P.and Lanka, E. (1989) TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer region. *J. Biol. Chem.* **264**: 11989-11994.

Ziegelin, G., Pansegrau, W., Lurz, R., Lanka, E. (1991) DNA binding properties of the gene products of the specificity deteminants *traJ* and *traK* of plasmid RP4. EMBO Workshop on bacterial conjugation systems.pp 92-94.